U.PORTO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO
U.PORTO
FMUP FACULDADE DE MEDICINA
UNIVERSIDADE DO PORTO
U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Human brain development and evolution: Insights from gene expression

André Miguel Moura da Costa e Sousa

Tese de doutoramento em Biologia Básica e Aplicada

2013

André Miguel Moura da Costa e Sousa

Human brain development and evolution: Insights from gene expression

Tese de Candidatura ao grau de Doutor em Biologia Básica e Aplicada submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador – Nenad Šestan

Categoria – Full Professor

Afiliação – Yale University School of Medicine

Co-orientador – José Manuel Pereira Dias Castro Lopes

Categoria – Professor Catedrático

Afiliação – Faculdade de Medicina da Universidade do Porto

"Not all those who wander are lost"

John R. R. Tolkien

ABSTRACT

Discovering the basis for the unique properties of the human brain is a major goal of modern science. Yet, many fundamental gaps remain in our knowledge of the human brain development and evolution.

In this study we analyzed the transcriptome from sixteen human brain regions, comprising the cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus, and eleven areas of the neocortex, throughout the entire prenatal development and postnatal life. Approximately 86% of genes were found to be expressed, and over 90% of these were differentially expressed across regions and/or time. Genes were organized into functionally distinct co-expression networks. Furthermore, sex differences were present in both gene expression and exon usage. We also demonstrate that these results can be used to profile trajectories of genes associated with several neurobiological themes, such as developmental processes, cell types, neurotransmitter systems, autism, and schizophrenia.

In the second part of this study, we analyzed the temporal dynamics and laterality of gene expression across the eleven neocortical areas. We found that inter-areal transcriptional differences exhibit a temporal hourglass pattern, characterized by more expression differences prenatally and from adolescence onward, and almost no differences in infancy and childhood. Further analyses revealed distinct gradients and areal-restricted expression. We also found that global gene expression trajectories of each area become increasingly synchronized, especially after birth. Furthermore, analyses of gene expression between corresponding left and right areas revealed global symmetry throughout the fetal development and postnatal lifespan.

Finally, we analyzed the gene expression from human, chimpanzee, and rhesus macaque adult brains, using the aforementioned sixteen regions. We found 6,389 genes were differentially expressed among species, with 3,154 specifically up- or downregulated in humans. Furthermore, a multi-regional approach allowed us to examine intra-species differentially expressed genes among all regions. In addition, we found region-specific gene co-expression modules that are conserved across all three species, as well as species-specific transcription modules. We also found several genes involved in the dopamine synthesis pathway were upregulated in humans, especially in the striatum. Lastly, we analyzed the co-expression patterns of neurotransmitter receptors systems and found that the glutamatergic and GABAergic systems receptors were the most conserved among species, while serotoninergic and cholinergic systems were the least conserved.

This study provides a comprehensive dataset on the spatiotemporal human brain transcriptome and new insights into the transcriptional foundations of human neurodevelopment, as well as new insights into species-specific gene expression patterns.

RESUMO

Um dos grandes desafios da ciência moderna consiste em descobrir a base biológica das propriedades únicas do cérebro humano. No entanto, ainda existem muitas lacunas no nosso conhecimento sobre o seu desenvolvimento e evolução.

Na primeira parte deste estudo, o transcriptoma de dezasseis regiões do cérebro, incluindo o córtex cerebelar, o núcleo dorsomedial do tálamo, o estriado, a amígdala, o hipocampo e onze áreas do neocórtex, foi analisado ao longo de todo o desenvolvimento pré e pós-natal. Verificou-se que cerca de 86% dos genes são expressos e que mais de 90% destes estão diferencialmente regulados ao nível do gene ou exão em todas as regiões e/ou tempo. A grande maioria dos genes foi organizada em diferentes redes funcionais de co-expressão, e também foram encontradas diferenças entre sexos ao nível da expressão génica. Adicionalmente, confirmou-se que estes resultados podem ser utilizados para perfilar trajectórias de genes associados com os processos de neurodesenvolvimento, de diferentes tipos de células, sistemas neurotransmissores e também algumas doenças neurológicas, como o autismo e a esquizofrenia.

A dinâmica temporal e lateralidade da expressão génica nas onze áreas neocorticais foram analisadas na segunda parte do estudo. Observou-se que as diferenças entre áreas, ao nível do transcriptoma, exibem um padrão temporal em forma de ampulheta, e que este é caracterizado por grandes diferenças tanto na expressão pré-natal como a partir da adolescência, apresentando um hiato na infância. Também se verificou que existem gradientes distintos de expressão, incluindo alguns com diferenças entre humano e macaco Rhesus, bem como transcrição restrita a determinadas áreas. A análise global das trajetórias de expressão génica de cada área também revelou que a maturação se torna cada vez mais sincronizada, especialmente durante o desenvolvimento pós-natal. Por último, a análise da expressão génica entre os hemisférios direito e esquerdo revelou que existe uma simetria global durante o desenvolvimento fetal e pós-natal.

Na última parte deste estudo, foi analisada a expressão génica das dezasseis regiões acima referidas em cérebros adultos de humanos, chimpanzés e macacos Rhesus. Detetaram-se 6389 genes diferencialmente expressos entre as espécies, com 3154 especificamente sobre ou subexpressos em humanos. Uma abordagem multi-regional também permitiu analisar genes diferencialmente expressos entre todas as regiões, em cada espécie, e comparar como as áreas se correlacionam de modo diferente entre as espécies analisadas. Além disso, verificou-se que existem módulos de co-expressão de genes que são conservados nas três espécies, bem como módulos de

transcrição específicos para cada espécie. Também se demonstrou que vários genes envolvidos na via de síntese da dopamina se encontram sobrexpressos em humanos, especialmente no estriado. Finalmente, analisaram-se os padrões de co-expressão de sistemas de receptores de neurotransmissores e verificou-se que os sistemas de receptores glutamatérgico e GABAérgico eram os mais conservados entre as espécies, ao passo que os sistemas serotoninérgico e colinérgico eram os menos conservados.

Este estudo fornece um conjunto abrangente de novos dados sobre as fundações biológicas da transcrição génica durante o neurodesenvolvimento humano e também sobre diferentes padrões de expressão de genes específicos da espécie humana.

ACKONWLEDGEMENTS

I would like to acknowledge my supervisor, Professor Nenad Šestan, for the opportunity to work on this project, for his mentorship and guidance through the research process, and for granting me the freedom of planning and doing the experiments that culminated in the writing of this thesis. I would like to extend the ackowledgement to all the Šestan Lab for creating an amazing environment for research. I have to acknowledge Mihovil Pletikos, Goran Sedmak, Hyo Jung Kang, and Yuka Imamura Kawasawa for their technical advice and help, and Ying Zhu, Feng Cheng, Mingfeng Li, and Xuming Xu for their help and for sharing their knowledge on data analyses. I also thank Kyle Meyer for the thoughtful discussions and for being always ready to help in any theoretical or pratical issue.

I would like to thank the GABBA Program, especially the coordinators of the 11th edition: Professor José Castro Lopes and Professor Fátima Carneiro. Thank you for your help during the courses and the laboratory selection process. I would like to especially ackowledge Professor José Castro Lopes, my co-supervisor for helpful discussions and guidance, and Professor Alexandre do Carmo, the Program's director, for all the help in dealing with University of Porto bureaucraticy. I have to give a very special ackowledgement to Professor Maria de Sousa for teaching me that "Life as a PhD student should only be fun".

Finally, I would like to deeply thank my family and friends for all the support during these five years.

This work would not have been possible without the generous support from the Portuguese Foundation for Science and Technology through the fellowship SFRH/BD/33255/2007.

TABLE OF CONTENTS

Abstract v
Resumo vii
Acknowledgements ix
Table of contents xi
List of Abbreviations xiii

CHAPTER I - Introduction 1
1. Human brain development 2

2. Interspecies phenotypic differences 6
2.1. Changes in timing 6
2.2. Brain size differences 7
2.3. Changes in microstructure and connectivity 9

3. Gene expression in the human brain 11

4. Transcriptome evolution 12
4.1. Developmental differences in the transcriptome 13

5. Aims 15

CHAPTER II - Results 19
1. Spatio-temporal transcriptome of the human brain 21
1.1. Abstract 22
1.2. Introduction 23
1.3. Results 23
1.4. Conclusions 30
1.5. Methods summary 32
1.6. Acknowledgements 32
1.7. Tables 33
1.8. Figures 34
1.9. Supplementary Information 40
1.10. References 109

2. Temporal dynamics and bilaterality of human neocortical areal transcription 115
2.1. Abstract 116
2.2. Introduction 117
2.3. Results 118
2.4. Discussion 124
2.5. Methods summary 126
2.6. Acknowledgements 126

2.7. Figures 127
2.8. Supplementary Information 132
2.9. References 159

3. Multi-regional analysis of gene expression evolution in the human brain 161
3.1. Abstract 162
3.2. Introduction 163
3.3. Results 163
3.4. Conclusions 168
3.5. Methods summary 169
3.6. Acknowledgements 169
3.7. Tables 170
3.8. Figures 171
3.9. Supplementary Information 177
3.10. References 208

CHAPTER III - General discussion, conclusions, and future directions 217

CHAPTER IV - References 227

CHAPTER V - Publications 259

LIST OF ABBREVIATIONS

A1C	primary auditory cortex
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
AMY	amygdala
ANOVA	analysis of variance
ASD	autism spectrum disorder
BA	Brodmann area
CALB1	calbindin
CALB2	calretinin
CBC	cerebellar cortex
CGE	caudal ganglionic eminence
CNV	copy number variation
CP	cortical plate
CV	coefficient of variation
DAB	3, 3'-diaminobenzidine
DABG	detection above background
DCX	doublecortin
DDC	L-dihydroxyphenylalanine decarboxylase
ddPCR	droplet digital polymerase chain reaction
DEU	differential exon usage
DEX	differential expression/differentially expressed
DFC	dorsolateral prefrontal cortex
DIE	diencephalon
DS	dissection score
DTH	dorsal thalamus
eQTL	expression quantitative trait loci
FC	frontal cortex
FDR	false discovery rate
Fig.	figure
GABA	γ-aminobutiric acid
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GLM	generalized linear model
GO	gene ontology
h	*Homo sapiens*
HIP	hippocampus
HWE	Hardy-Weinberg-Equilibrium
IHC	immunohistochemistry
IPC	posterior inferior parietal cortex
ITC	inferior temporal cortex
L	layer
L-DOPA	L-dihydroxyphenylalanine
LGE	lateral ganglionic eminence
LOWESS	locally weighted scatterplot smoothing
m	*Macaca mulatta*
M	module

M1C	primary motor cortex
Ma	Million years ago
MD	mediodorsal nucleus of the thalamus
MDS	multidimensional scaling
MFC	medial prefrontal cortex
MGE	medial ganglionic eminence
MSC	motor and parietal somatosensory cortices
MZ	marginal zone
NADPH	nicotinamide adenine dinucleotide phosphate
NADPH-d	nicotinamide adenine dinucleotide phosphate diaphorase
NCX	neocortex
NHP	non-human primate
NMDA	N-methyl-D-aspartate
OFC	orbital prefrontal cortex
OP	occipital cortex
p	*Pan troglodytes*
PAH	phenylalanine hydroxylase
PBS	phosphate buffered saline
PC	parietal cortex
PC	principal component
PCA	principal component analysis
PCC	Pearson correlation coefficient
PCR	polymerase chain reaction
PCW	weeks post-conception
PMI	postmortem interval
PVALB	parvalbumin
QC	quality control
qPCR	quantitative real-time polymerase chain reaction
r^2	coefficient of determination
RIN	RNA inegrity number
RNA-seq	RNA-sequencing
RPKM	read per kilobase per millions of mapped reads
S1C	primary somatosensory cortex
SATB2	special AT-rich sequence-binding protein 2
sd	standard deviation
SGL	subpial granular layer
SNP	single nucleotide polymorphism
SP	subplate
STC	superior temporal cortex
STR	striatum
TBR1	T-box brain 1
TC	temporal cortex
TES	transcription end site
TH	tyrosine hydroxylase
TSS	transcription start site
URL	upper (rostral) rhombic lip

UTR	untranslated region
UV	ultraviolet
V1C	primary visual cortex
VF	ventral forebrain
VFC	ventrolateral prefrontal cortex
VIP	vasoactive intestinal peptide
WGCNA	weighted gene co-expression network analysis
WM	white matter
Y	years

CHAPTER I

INTRODUCTION

Human beings possess incredible cognitive and sensorimotor capabilities. Every day we communicate and organize thoughts using language, form and retrieve memories, and act in a creative way. This level of cognition, creativity, and sensorimotor skill is present only in humans. Even our evolutionary closest living relatives, the great apes, lack most of these capabilities. Thus, it is of no surprise that we have been fascinated with uncovering how our brain is capable of producing such skills (Jerison 1973, Kaas and Preuss 2007, Gazzaniga 2008).

Most of our current understanding of the molecular mechanisms governing brain development and function are derived from experimental studies of a handful of model organisms, such as the fruit fly, frog, mouse, and rat. While the use of these model organisms has revolutionized biomedical sciences, these systems cannot fully reproduce features of human brain development and dysfunction, particularly those affecting the most distinctly human aspects of cognition and behavior. Comparing humans to non-human primates (NHPs) is necessary for defining uniquely human features in both healthy and diseased states, but this comparison is often difficult to do in a systematic manner. Advances in non-invasive technologies such as diffusion tensor imaging (DTI) are expanding the range of comparison studies that are possible. As the number of large-scale and precise comparisons that are feasible increases, a major challenge will be discerning how changes at the genomic level influence different phenotypes. Mapping molecular evolution to phenotypic evolution is a daunting task in model organisms, and accomplishing this in humans and NHPs presents further obstacles. Ethical considerations as well as technological limitations prevent thorough experimentation, and analysis of transcription and protein expression will be always confined to postmortem samples.

Although discovering the genetic basis and molecular mechanisms that contribute to the differences concerning human behavior and cognitive abilities is certainly of great interest, the current feasibility of this remains a challenging task. While still challenging, determining the changes that underlie differential brain organization and development arguably provides an entry point, one that can be built upon to reach the most distinctly human aspects of cognition and behavior.

1. Human Brain Development

Human brain development is a highly dynamic process during which different regions undergo distinct maturational changes (Retzius 1896, His 1904, Sidman and Rakic 1973, Poliakow 1979, Kostovic and Judas 2002, Kang, Kawasawa et al. 2011). Moreover, transient brain structures arise and disappear during specific developmental periods. Thus, in order to have a comprehensive analysis of the human brain development, it is crucial to include multiple regions and major developmental time points, allowing the identification of temporally- and spatially-specific changes.

At the present, well-defined morphological staging is limited to the embryonic development (O'Rahilly and Müller 2006), and no fully satisfactory staging system has yet been devised for the fetal and early postnatal development. In this study, we have specified time periods based mainly on the timing and progression of major neurodevelopmental events in the cerebral cortex, a structure that is central to the highest cognitive functions in humans and is one of the most studied human brain structures. Our system took into consideration previous developmental classifications of the human brain (Poliakov 1949, Sidman and Rakic 1973, Sidman and Rakic 1982, O'Rahilly and Müller 2006, Kostovic and Judas 2009, Judas 2010).

Human brain development can be broadly divided into two periods: prenatal and postnatal. Although prenatal development is relatively short in comparison to postnatal development, it is highly dynamic. Therefore, we divided prenatal development into seven distinct periods to facilitate higher temporal resolution. For postnatal developmental periods, we considered cognitive, motor, social, and emotional milestones outlined by the Department of Human Health and Services (http://www.cdc.gov/ncbddd/child/). As previously elaborated (O'Rahilly and Müller 2006), prenatal age is with respect to fertilization by definition. For prenatal age we used postconceptional weeks (PCW) to include both proper terms used in current literature: postfertilizational (timed from the fertilization of the ovum) and postovulatory (timed from the ovulation). Although there is a small difference in the timing of these events, for the majority of purposes these three terms (postconceptional, postfertilizational and postovulatory) could be used synonymously. The following is a brief description of the periods as defined in this study:

Period 1 (Embryonic development, 4≤ Age <8 PCW) corresponds to late embryonic development defined by the first lamination of the cerebral wall (i.e., ventricular zone, intermediate zone, marginal zone). Early embryonic processes (e.g., formation of the neural tube, closure of the neuropores, and formation of the primary and secondary brain vesicles) are completed in this period. This period is marked by extensive proliferation and the initiation of neurogenesis. The first axons invade the cerebral wall during this period.

Period 2 (Early fetal development, 8≤ Age <10 PCW) is characterized by the appearance and subsequent primary consolidation of the cortical plate. Deep layer neurons are generated and begin to radially migrate to their proper position in the cortical plate. The secondary proliferative zone, the subventricular zone, appears around 8 PCW. The internal capsule and anterior commissure begin to appear. Major neuroanatomical fetal landmarks are readily recognized during this period, including the ventricular zone, subventricular zone, intermediate zone, presubplate zone (i.e., precursor of future subplate proper), cortical plate, marginal zone, and subpial granular layer can be distinguished in the cerebral wall.

Period 3 (Early fetal development, 10≤ Age <13 PCW) is characterized by the presence of the bilaminate cortical plate. Namely, the deep part of the cortical plate itself becomes delaminated and together with a thin, cell-sparse band of tissue (described as the presubplate in the previous period) represents the subplate in formation. The first synapses are formed in the marginal zone above the cortical plate and in the presubplate zone below the cortical plate, whereas the cortical plate remains free of synapses. Tangential migration of GABAergic interneurons can be observed. Proliferation and migration of neurons are the two main histogenetic processes that occur during this period. The corpus callosum can be identified during this period. Afferent projections start to invade the cortical anlage and are primarily monoaminergic. The first sulci start to appear in this period (i.e., lateral and callosal sulcus).

Period 4 (Early mid-fetal development, 13≤ Age <16 PCW) is characterized by the secondary consolidation of the cortical plate concomitantly with the formation of a large, synapse-rich subplate zone. The first cortical neurons show signs of morphological differentiation. After its formation and during subsequent mid-fetal periods, the subplate zone serves as a waiting compartment for afferent axons from several subcortical structures. Proliferation and migration of neurons and ingrowth of afferent axons are the major histogenetic events occurring during this period. At this period and subsequent prenatal periods, typical fetal lamination pattern of the cerebral wall can be observed. The cerebral wall can be divided into six major architectonic compartments or fetal zones: ventricular zone, subventricular zone, intermediate zone, subplate zone, cortical plate, and marginal zone, which contains a transient subpial granular layer.

Period 5 (Early mid-fetal development, 16≤ Age <19 PCW) is characterized by an increase in the size of the subplate zone and the overall thickness of the cerebral wall. Ingrowth of thalamocortical axons into the subplate and migration of upper cortical layer projection neurons are major histogenetic events. The parieto-occipital, cingulate, and calcarine sulci appear in this period.

Period 6 (Late mid-fetal development, 19≤ Age <24 PCW) is characterized by the peak of subplate zone thickness and development, and by the massive ingrowth of afferent axons from the subplate zone into the cortical plate. Therefore, the first synapses can be observed in the neocortical plate. Neurogenesis ceases in the pallial wall but still continues in the subpallial ganglionic eminences during this period. The central sulcus, superior temporal sulcus, collateral sulcus, superior temporal gyrus, and parahippocampal gyrus can be identified.

Period 7 (Late fetal development, 24≤ Age <38 PCW) is characterized by the transformation of the typical fetal lamination pattern into an adult-like lamination pattern of the cerebral wall. Resolution of the subplate zone starts and the Brodmann's "six-layered ontogenetic Grundtypus" (i.e., the fetal equivalent of future layers 2–6) appears within the cortical plate during this period. Cytoarchitectonic regional and areal differentiation of the cortical plate/cortex is an important event that occurs during this period. Neuronal differentiation, ingrowth of thalamocortical axons, and gliogenesis are major histogenetic processes observed during this period. Synaptogenesis continues primarily in the cortex. Myelination of select cortical axon projections starts in this period. The ventricular zone gradually thins until it appears as a single layer of cuboidal/columnar cells. Majority of cortical gyri and sulci appear during this period.

Period 8 (Neonatal and early infancy, birth≤ Age <6 postnatal months) is characterized by reorganization of long afferent and corticocortical axons, transformation and maturation of cortical layers (especially layer 5 and 6) from a fetal to an adult-like pattern, and rapid synaptogenesis and spinogenesis. The remnants of the subplate zone, although resolving, are still present below layer 6 and thus form a transition from immature cortex to the developing gyral white matter. Specific motor (e.g., grasping, raising of the head, stretching of the legs and kicking) and sensory (e.g., object following, head turning toward sound) skills appear during this period.

Period 9 (Late infancy, 6≤ Age <12 postnatal months) is characterized by further development of motor skills (e.g., sitting with and without support), sensory skills (e.g., development of color vision), and cognitive skills (e.g., exploration with hand and mouth, response to one's own name). Resolution of the subplate zone is completed and neurons that survive resolution of the subplate zone are incorporated into the subcortical (gyral) white matter as interstitial neurons.

Period 10 (Early childhood, 1≤ Age <6 years) is characterized mainly by reorganization and maturation of local circuits, and the peak of synaptogenesis. During this period rapid development of motor skills (e.g., walking, pincer grasp, fine movement control), social

and emotional skills (e.g., development of imitation, self-awareness, independence), and cognitive skills (e.g., thinking, mathematical abilities, language) is observed.

Period 11 (Middle and late childhood, 6≤ Age <12 years) is characterized by further cognitive development and refinement of neural circuits.

Period 12 (Adolescence, 12≤ Age <20 years) is characterized by sexual maturation and the appearance of adult-like connectivity pattern.

Period 13 (Young adulthood, 20≤ Age <40 years) is characterized by the end of maturation processes in the brain (e.g., myelination ends in the first part of this period) and appearance of an adult-pattern of brain functions.

Period 14 (Middle adulthood, 40≤ Age <60 years) is characterized by an adult-like pattern of brain functions and the beginning of aging processes.

Period 15 (Late adulthood, 60 years +) is characterized by the progression of aging processes.

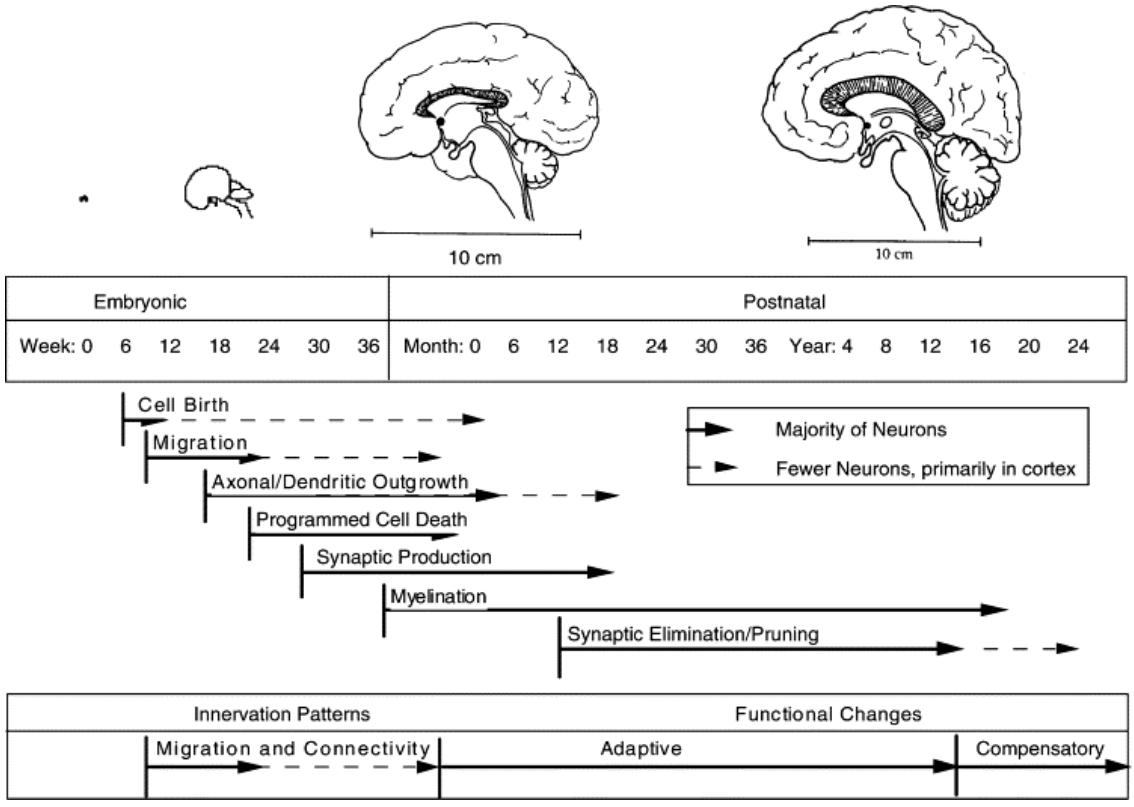


Figure 1. Timeline of major events in human brain development.

This diagram represents human brain development beginning with conception, and proceeding with cell generation, neuronal migration, axonal/dendritic outgrowth, synaptogenesis, myelination, and synaptic pruning. From (Andersen 2003).

2. Interspecies phenotypic differences

2.1. Changes in timing

The differences in structure between human and NHP brains are (at least mostly) due to changes in brain development, but the underlying molecular and cellular mechanisms driving these changes are not well understood. One of the main advances was the discovery of a correlation between the timing and duration of neurogenesis, corticogenesis in particular, and the size of brain structures, especially the neocortex (Finlay and Darlington 1995). It was found that human neurogenesis is protracted, leading to an increase in progenitor cells that give rise to the neocortex. As an example, upper layer neurons, which are the last postmitotic neurons to migrate to their final position in the laminar neocortex (Angevine and Sidman 1961, Rakic 1974), have an extraordinary expansion in numbers in primates, especially in humans, when compared to other mammals (Charvet, Striedter et al. 2011). This further corroborates that the delayed and prolonged neurogenesis is a major factor in creating the interspecies differences in brain size we observe in postnatal brains. However, this delay and prolongation is not restricted to neurogenesis. Humans have a very prolonged and protracted development (Huttenlocher and Dabholkar 1997, Miller, Duka et al. 2012, Petanjek and Kostovic 2012). Furthermore, there are several studies, from gene expression to morphology, that show that we have several neotenous traits (Gould 1977, Somel, Guo et al. 2010, Miller, Duka et al. 2012). Like other NHPs, humans require parental support during childhood, and because this developmental stage is longer in humans, it is one of the key factors for our social specializations (Bogin 1988).

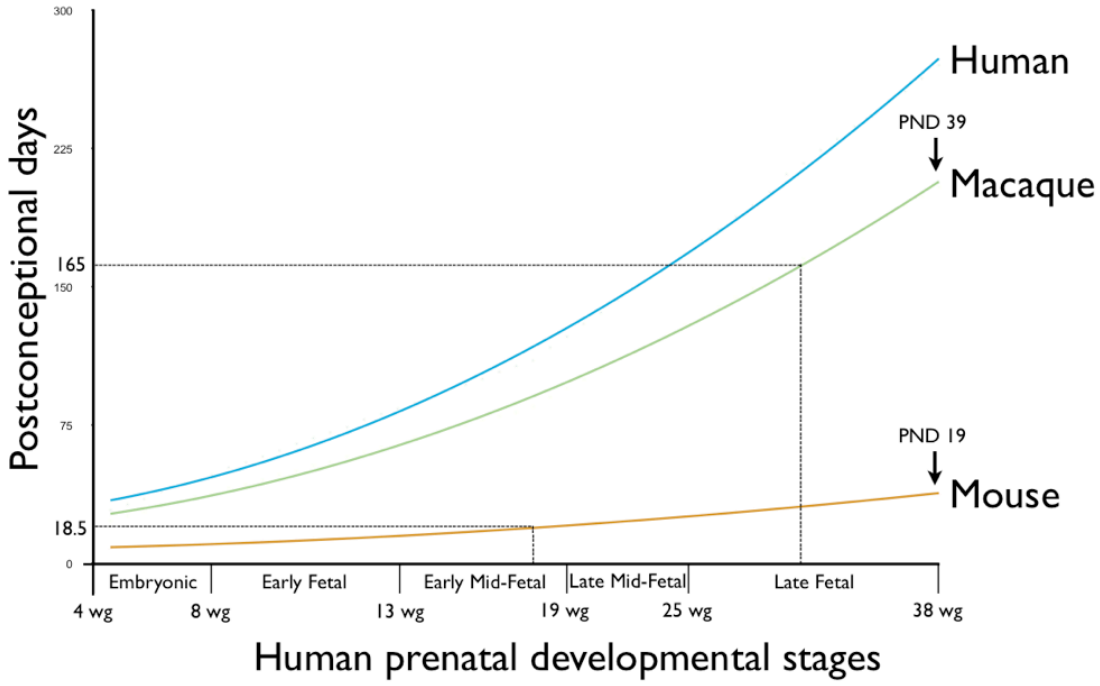


Figure 2. Comparative brain developmental trajectories of human, macaque, and mouse. This figure shows the heterochronicity in development between species. Mouse prenatal development lasts for around 18.5 days, while macaque prenatal development completes in 165 days. When mice are born, the human brain is still in its early mid-fetal development. When macaques are born, humans are already in late fetal development. The human brain development time at birth corresponds to 19 postnatal days in mice and 39 postnatal days in macaque, approximately. This figure was plotted based on (Clancy, Kersh et al. 2007). The developmental points of the 3 species were normalized to the rat development based on an extensive list of morphological features.

2.2. Brain size differences

The average adult human brain weighs approximately 1,300 g (DeFelipe 2011), which triples the average size of a chimpanzee brain (Sherwood et al., 2008). As the absolute size increases, the brain is not expected to scale uniformly. For example, the neocortex tends to make up a greater proportion of larger brains (Frahm, Stephan et al. 1982) and neocortical white matter seems to increase disproportionately compared to gray matter (Hofman 1985, Rilling and Insel 1999, Zhang and Sejnowski 2000). By taking these differences in scaling into account, it can be asked whether areas in the human brain are the size that would be expected for the overall size of the human brain. Association areas are of particular interest for cognitive functions. When compared to other primates, it seems the size of human association areas are what is expected for the size of the human

brain (Sherwood, Subiaul et al. 2008). Likewise, the size of the frontal cortex seems no larger than expected compared to the size of the rest of the brain (Semendeferi, Damasio et al. 1997, Semendeferi and Damasio 2000, Semendeferi, Lu et al. 2002, Bush and Allman 2004). Within the frontal cortex, Schoenemann (2005) and colleagues provide evidence the prefrontal cortex is expanded, but this finding may have resulted from the methodology of the study (Sherwood, Holloway et al. 2005). Area 10 of the prefrontal cortex has also been reported to be relatively larger in humans (Semendeferi, Armstrong et al. 2001), but the degree of this increase has been questioned (Holloway 2002). Furthermore, the cerebellar hemispheres also appear to be larger than expected (Rilling and Insel 1998, MacLeod, Zilles et al. 2003).

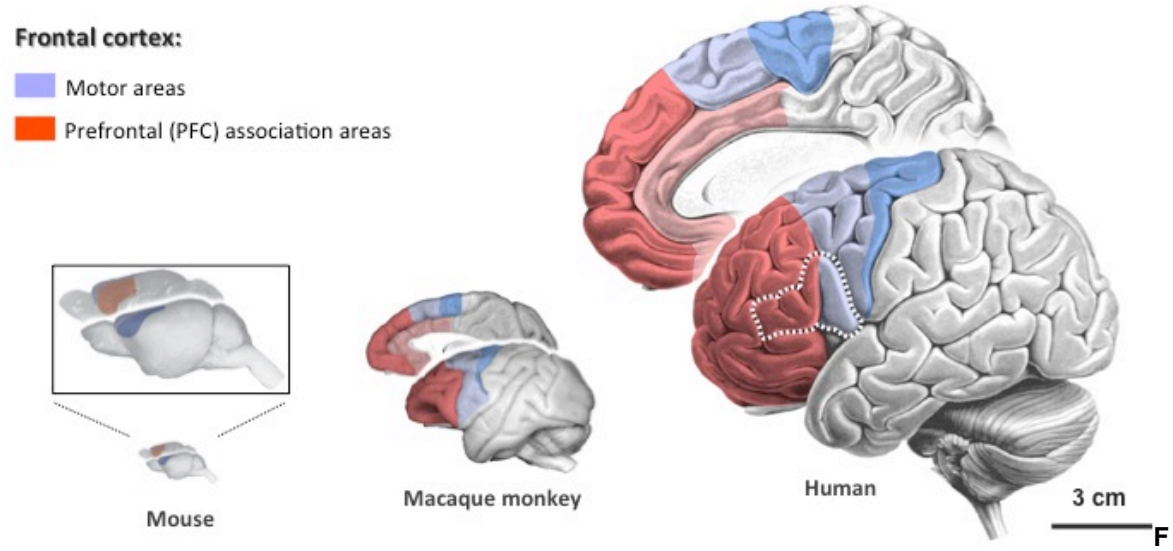


Figure 3. Comparison of mouse, macaque monkey, and human adult brains. Human brain is remarkably larger in terms of absolute size. Although motor (blue) and pre-motor (light blue) areas are conserved in terms of topography and relative size, prefrontal association areas, temporal lobe, and cerebellar hemispheres appear to be enlarged in humans, and restricted to the medial prefrontal cortex in mouse (Preuss, 1995).

While currently there is not strong evidence that some areas have increased relative to brain size in the human lineage, interpreting these results is difficult. While allometric considerations provide an expectation (and humans fall within this range), it does not follow that these changes are behaviorally irrelevant. After all, absolute brain size, not relative measures, seems to best predict cognitive abilities across primates (Deaner, Isler et al. 2007). Moreover, consideration of the number and morphology of both neurons and glia that occurred in human lineage should be taken into account (Elston, Benavides-Piccione et al. 2006, Herculano-Houzel 2011, Bianchi, Stimpson et al. 2012).

2.3. Changes in microstructure and connectivity

Do humans have any structures or areas not present in other primates? Despite the large expansion of the cortex with respect to absolute size, current evidence does not support that there are any cortical areas unique to humans (Preuss 2011). This includes human peryslvian areas involved in speech and language, which appear to be present in at least some NHPs (Preuss and Goldman-Rakic 1991, Grefkes and Fink 2005, Petrides, Cadoret et al. 2005). Likewise, histological comparisons have not identified human-specific changes in organization, with the exception of layer 4a in the human visual cortex (Preuss and Coleman 2002).

In contrast, there have been several studies supporting human-specific connectivity patterns. One example is the arcuate fasciculus, a pathway that connects temporal-parietal and frontal cortical areas. It has been shown that alterations of the arcuate fasciculus and associated cortical areas result in conduction aphasia (Anderson, Gilmore et al. 1999), indicating the importance of these fibers for human language. DTI analysis of this tract in humans, chimpanzees, and macaques revealed that temporal projections present in humans seem to be greatly reduced in chimpanzees and absent in macaques (Rilling, Glasser et al. 2008).

It has also been proposed that there is an increased connectivity of associative areas in the human neocortex, including long cortico-cortical connections and short U-shaped fibers (Catani, Dell'acqua et al. 2012). The latter connect neighboring areas and are thought to generate more elaborate gyrification (Van Essen 1997). In a recent study (Spocter, Hopkins et al. 2012), the neuropil fraction (used as a proxy of the total connectivity, assuming that more neuropil implicates more connectivity) of six different neocortical areas was quantified in primary and associative areas of humans and chimpanzees. They found that human associative areas have higher fraction of neuropil, especially the prefrontal areas, while chimpanzee associative areas did not show a significant difference for neuropil levels. This evidence that neurons are more separated in certain areas of the neocortex was also found in associative areas of the frontal (Buxhoeveden, Switala et al. 2001, Schenker, Buxhoeveden et al. 2008) and the temporal (Semendeferi, Teffer et al. 2011) cortices. These results reinforce the idea that humans have an increased connectivity, especially in the prefrontal cortex (PFC), a region involved in higher cognitive functions.

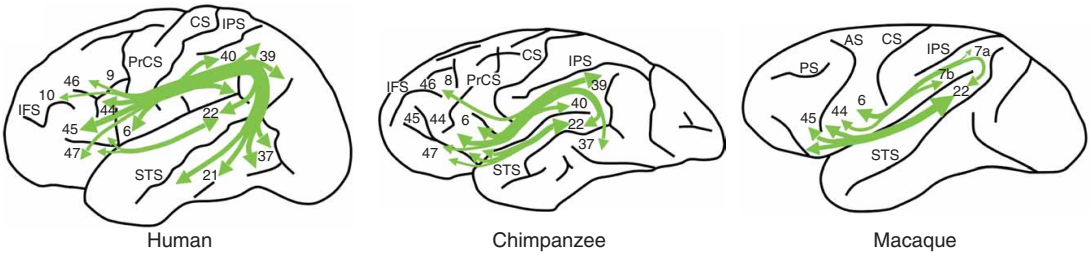


Figure 4. Arcuate fasciculus differences among primate species.
Scheme showing the organization of the arcuate fasciculus in human, chimpanzee, and macaque brains. Remarkably, the human frontal cortex, but not the chimpanzee or the macaque, of the left hemisphere is strongly connected via the arcuate fasciculus with the left medial and inferior temporal gyrus. From Rilling et al., 2008. (Rilling, Glasser et al. 2008)

It is very important to notice that many of the human specializations discussed above seem to be elaborations on specializations shared by all primates, especially the great apes. When compared to other mammals, primates also have larger brains, specific brain areas (especially prefrontal cortex and temporal cortex specializations), and particular developmental features (Preuss 2000, Sherwood, Subiaul et al. 2008).

3. Gene expression in the human brain

A very high percentage of the known protein-coding genes are expressed in the human brain (Colantuoni, Purcell et al. 2000, Myers, Gibbs et al. 2007). When compared with several other organs, the human brain has a distinct transcriptional profile (Roth, Hevezi et al. 2006), higher levels of gene expression, and the expressed transcripts are more diversified (Ramskold, Wang et al. 2009, de la Grange, Gratadou et al. 2010). There are several factors that may be responsible for this diversification of RNA populations, such as alternative splicing and the transcription of noncoding RNAs. It has been shown that there is an enrichment for splicing events (such as skipped exons, alternative poly(A) signal usage, and alternative promoter usage) in transcripts in the brain (Yeo, Holste et al. 2004, Wang, Scully et al. 2007). There is also evidence that both long noncoding RNAs and microRNAs are also enriched in the brain (Kuss and Chen 2008, Ponjavic, Oliver et al. 2009, Chodroff, Goodstadt et al. 2010). All these studies revealed very relevant features of the brain when compared with other tissues. But, since the brain is a very complex organ and it is composed of several histologically and functionally distinct structures, studies of gene expression on the multitude of regions in the human brain may reveal crucial information on how the brain is constructed. As an example, studies on the human brain development across several regions reported that the cerebellum has a very distinct profile in terms of gene expression (Roth et al., 2006; Lockhart and Barlow, 2001). Importantly, it is also important to study these regions through development, because some of the patterns that govern complex processes during regional specification may be transient.

Most gene expression studies of the developing brain have been restricted to rodents (Zapala, Hovatta et al. 2005, Semeralul, Boutros et al. 2006). From those performed in humans, most have included a relatively small number of samples and are predominantly focused on few regions, or developmental time points (Sun, Patoine et al. 2005, Abrahams, Tentler et al. 2007, Johnson, Kawasawa et al. 2009, Ip, Wappler et al. 2010). Also, the fact that significant asymmetry in gene expression between the left and right hemispheres was found during early fetal development (Sun, Patoine et al. 2005), but was not detected during midfetal development (Johnson, Kawasawa et al. 2009, Lambert, Lambot et al. 2011), reinforces that it is necessary to extend the study of gene expression to more developmental stages, including fetal and postnatal development.

4. Transcriptome evolution

It is often suggested that differences in expression of genes, rather than the makeup of the genes themselves, have been the major drivers of phenotypic evolution. One motivation for this claim is the argument, which is proposed on the basis of exclusion, that the protein-coding differences are too small to account for this (King and Wilson 1975). Other arguments for the importance of regulatory evolution are rooted in the idea of circumventing the increased selective pressure on genes with pleiotropic effects (Blekhman, Oshlack et al. 2008). Similar to how a duplication of a gene relaxes selective pressure, modifications of regulatory elements allow tissue- and time-specific changes in gene expression levels. This increases the probability that a regulatory mutation that confers a selective advantage will reach fixation because it does not need to outweigh any disadvantages caused by an increase in expression in all tissues. Perhaps more importantly, most human-specific features, including those discussed above, are quantitative rather than qualitative (that is, they are present but to a different degree in other species). Thus, changes in developmental processes and timing via regulatory evolution, are likely to be heavily involved (Carroll 2003).

Global studies comparing transcription in humans and NHPs provide a method for determining human-specific gene expression. Even though differences in gene expression during development are crucial, most studies on transcriptome evolution were done in adult specimens due to the scarcity of developmental tissue in good working conditions, especially within great apes. A handful of these studies reported evidence that gene expression in the human brain has diverged more from other primate species than gene expression of other tissues examined (Enard, Khaitovich et al. 2002) and that there was a bias for upregulated expression in the human brain that was not observed in the other tissues sampled (Cáceres, Lachuer et al. 2003, Gu and Gu 2003, Khaitovich, Muetzel et al. 2004). However, other studies did not find higher divergence in brain gene expression (Hsieh, Chu et al. 2003) nor a bias for upregulated expression in the human brain (Uddin, Wildman et al. 2004). Some other works focused on differences in metabolic genes, especially in aerobic metabolism (Uddin, Wildman et al. 2004, Babbitt, Fedrigo et al. 2010), groups of human-specific co-expressed genes (Oldham, Horvath et al. 2006, Konopka, Friedrich et al. 2012), and also noncoding RNAs, which were surprisingly conserved in terms of expression, suggesting that they may have a conserved functional role (Babbitt, Fedrigo et al. 2010).

Most of these studies were carried out with microarrays, which have inherent limitations. For interspecies studies, reliable expression detection is limited to genes that are identical, as sequence mismatches decrease hybridization reliability (Preuss, Cáceres

et al. 2004). Thus, these analyses are confined to a subset of genes whose sequences are very conserved. Also, arrays are only able to detect expression within pre-designed probe sets, preventing discovery of unknown transcripts. Furthermore, lowly expressed genes are difficult to detect at confident levels because they have to be significantly detected above the background, while highly expressed genes are also hard for differential expression calculation, as the probes reach saturation for the binding RNA, masking the real expression of these genes (Yang, Dudoit et al. 2001, Hekstra, Taussig et al. 2003). These problems were partially reduced with the introduction of next generation sequencing. Using RNA-seq, it is possible to uncover new transcripts, since it is not based on pre-designed probes. Also, the detection and quantification of gene expression does not depend on hybridization rates, which facilitates the detection of lowly expressed genes as well as finding differences in expression for highly expressed genes. However, RNA-seq also has limitations, especially with respect to the quality of different genomes annotations. The relatively better annotation of the human genome makes it difficult to analyze interspecies expression data without having human-biased results. The improving annotation of NHP genomes will eventually resolve this problem (Liu, Cheng et al. 2010).
Another problem when comparing transcriptomes among primate species is the differences in environment, such as diet. These differences should be as small as possible, in order to confidently state that the differences observed result from evolutionary changes and not environmental differences. This is an impossible problem to circumvent due to both tissue scarcity and ethical concerns that prevent designing controlled experiments.

4.1. Developmental differences at the transcriptome level

As discussed above, humans have developmental patterns and timings that differ from other species, even from other NHPs, with most of the regional differences in gene expression occurring prenatally and then decreasing over time (Colantuoni, Purcell et al. 2000, Somel, Guo et al. 2010, Colantuoni, Lipska et al. 2011, Kang, Kawasawa et al. 2011). It is also expected that crucial evolutionary differences at the transcriptome level will also occur during development. However, it is hard to compare species transcriptomes through development for technical reasons (it is extremely hard to harvest tissue from developing NHPs, especially apes) and, more importantly, because of developmental heterochronicity. What age in chimpanzees or macaques corresponds to a 1-year-old human? This is a hard question to answer. Barbara Finlay has done impressive and critical work for prenatal comparisons (http://translatingtime.net), based on several morphological data sets (Clancy, Kersh et al. 2007). Recently, a dynamic time warping algorithm was also developed (Yuan, Chen et al. 2011) in order to analyze time shifts in

gene expression based on time series data. However, aligning developmental events continues to be one of the hardest things to manage when performing a large-scale study on developmental differences in gene expression.

Despite the difficulties in obtaining developmental specimens, several studies have focused on gene expression across developing primate brains. These studies highlight groups of genes with a developmental time-shift and focus on neotenic features (Somel, Franz et al. 2009), miRNA regulation (Somel, Liu et al. 2011), and transcriptional factor regulation (Liu, Somel et al. 2012) of groups of co-expressed genes. These differences are more pronounced in the human PFC when compared to cerebellum. Although this might be very interesting, because the PFC is one of the last brain regions to mature and known to be involved in human-specific capabilities, the results are difficult to interpret because no other neocortical areas were examined. Human neocortical areas are very similar in terms of expression during postnatal development (Kang, Kawasawa et al. 2011), suggesting that the difference observed between the cerebellum and PFC might not be specific to the PFC.

The observation that the transcription factor MEF2A might be regulating a module of co-expressed genes involved in synaptogenesis (Liu, Somel et al. 2012) highlights the fact that alteration of the expression time course of a single transcription factor may drive differences in crucial developmental processes. Genes in this module are expressed later in human development than in other NHPs concurrent with the expression of MEF2A. These expression changes match up with the delay of synaptogenesis observed in humans (Rakic, Bourgeois et al. 1986, Huttenlocher and Dabholkar 1997).

Further studies are needed to corroborate these findings and to show evidence for mechanisms governing these differences. More importantly, it is of the utmost importance to extend transcriptional evolution studies to prenatal stages. This will most probably reveal critical differences on the transcriptional program that shape the brain.

5. Aims

The main goal of this work has been the comprehensive analysis of the human brain transcriptome, with special focus on the neocortical areas, during the entire lifespan. Another aim was to reveal human-specific transcription in the adult brain, by comparing it to homologous regions of the chimpanzee and macaque brains. To achieve this, two main platforms were employed: the Affymetrix Exon Microarray and the Illumina RNA sequencing technology. These provided the capability for analyzing the transcriptome at a high resolution, allowing a comprehensive analysis of exon-level expression data. This study has uncovered several new features of the gene expression spatio-temporal dynamics during human neurodevelopment, including regional specification, differences between sexes, maturational properties, and interhemispherical symmetry. Furthermore, this analysis revealed human-specific patterns of gene expression, including differences in neurotransmitter receptors. These data open new avenues for the research on human evolutionary specializations.

CHAPTER II

RESULTS

This section contains the results and their respective discussion and materials and methods. It is organized into three sub-chapters presented in journal article format:

1 - Spatio-temporal transcriptome of the human brain

This article was published in *Nature* (2011). Here, we describe several aspects of sixteen regions of human brain throughout development, adulthood, and aging. We reveal that the human brain transcriptome changes very dynamically during development, especially prenatally. We also show that the transcriptome is organized into distinct co-expression network modules, and that there is a sex-biased gene expression, even at the level of exon usage. Furthermore, we also profile developmental trajectories of genes involved in several neurobiological processes, like migration, differenciation, and synapse formation. Finally, we identified associations between single nucleotide polymorphism and gene expression. I contributed with tissue dissection, tissue sectioning, histological stainings and analysis, gene ontology analysis, selection and qRT-PCR validation of genes of special interest, and organizing the supplementary information. I also participated in the analyses of results. The manuscript was written by Nenad Šestan.

2 - Temporal Dynamics and Bilaterality of Human Neocortical Areal Transcription

This article was recently submitted to *Science* and is currently under revision. Here, we show that neocortical areas differences in gene expression decrease drastically perinatally and start to increase from adolescence on, exhibiting, therefore, an hourglass pattern. We also show evidence that some of the genes that have temporally regulated inter-areal differential expression also display divergence, both spatially and temporally, between human and monkey. These differences in gene expression patterns between species may have played a role in human evolution. Furthermore, using global transcriptional developmental trajectories, we investigated the transcriptional dynamics associated with areal maturation. Our findings suggest that, globally, there are no significant differences in maturational trajectories among neocortical areas, even though some neurodevelopmental processes, such myelination and synaptogenesis, show differences among human neocortical areas. Finally, we show that the neocortex is transcriptionally symmetric across the full course of fetal development and postnatal lifespan. I participated in all the experiments, analyzed the results, and wrote parts of the manuscript and supplementary information.

3 - Multi-regional analysis of gene expression evolution in the human brain

This article is currently under preparation for submission. Here, we describe several differences between human and other NHPs (chimpanzee and rhesus macaque) brain transcriptomes. We show that the species, and not the brain regions, contribute the most for the differences observed. We also show that the striatum is the region that shows more human-specific differential expression, and that the majority of the genes that have a human-specific pattern of expression are usually upregulated. Interestingly, the number of human-specific genes is higher than the number of chimpanzee-specific genes. On the other hand, the number of intra-species DEX genes is lower in humans, especially in the NCX. We also reveal several modules of co-expressed genes, including one with human-specific expression in the NCX, and another with human-specific expression in the STR, AMY, and HIP. Intriguingly, we found that the genes that code for enzymes involved in the anabolic pathway of dopamine are present in these modules. After a careful analysis, we found that both *TH* and *DDC* were upregulated in the human striatum, while *MAOA* and *COMT*, two of the enzymes that are involved in the catabolic pathway of dopamine, are not DEX, which means that, most probably, humans have more dopamine levels in the STR than chimpanzee and macaques. We also analyzed neurotransmitter receptors networks and found that while glutamatergic and GABAergic systems are very conserved among species, there are some differences in serotoninergic and cholinergic systems, which may contribute for human-specific behavior and cognition. In this study, I performed all the experiments, analyzed the results, and wrote the manuscript and parts of the supplementary information.

1. Spatio-temporal transcriptome of the human brain

Hyo Jung Kang[1*], Yuka Imamura Kawasawa[1*], Feng Cheng[1*], Ying Zhu[1*], Xuming Xu[1*], Mingfeng Li[1*], André M. M. Sousa[1,2], Mihovil Pletikos[1,3], Kyle A. Meyer[1], Goran Sedmak[1,3], Tobias Guennel[4], Yurae Shin[1], Matthew B. Johnson[1], Željka Krsnik[1], Simone Mayer[1,5], Sofia Fertuzinhos[1], Sheila Umlauf[6], Steven N. Lisgo[7], Alexander Vortmeyer[8], Daniel R. Weinberger[9], Shrikant Mane[6], Thomas M. Hyde[9,10], Anita Huttner[8], Mark Reimers[4], Joel E. Kleinman[9], Nenad Šestan[1]

[1]Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510, USA.
[2]Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical Sciences Institute, University of Porto, Porto, Portugal.
[3]Graduate Program in Neuroscience, Croatian Institute for Brain Research, University of Zagreb School of Medicine, Zagreb, Croatia.
[4]Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia 23298, USA. [5]International Max Planck Research School for Molecular Biology, Göttingen, Germany,
[6]Yale Center for Genome Analysis, Yale University School of Medicine, New Haven, Connecticut 06510, USA.
[7] Institute of Genetic Medicine, Newcastle University, International Centre for Life, Newcastle upon Tyne, NE1 3BZ, UK.
[8]Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06510, USA.
[9]Clinical Brain Disorders Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892, USA.
[10]The Lieber Institute for Brain Development, Johns Hopkins University Medical Center, Baltimore, Maryland 20892, USA.

* These authors contributed equally to this work.

Correspondence:
Nenad Sestan, MD, PhD
E-mail: nenad.sestan@yale.edu
Lab website: www.sestanlab.org

1.1. Abstract

Here we report the generation and analysis of genome-wide exon-level transcriptome data from 16 brain regions comprising the cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus, and 11 areas of the neocortex. The dataset was generated from 1,340 tissue samples collected from one or both hemispheres of 57 postmortem human brains, spanning from embryonic development to late adulthood and representing males and females of multiple ethnicities. We also performed genotyping of 2.5 million SNPs and assessed copy number variations for all donors. Approximately 86% of protein-coding genes were found to be expressed using stringent criteria, and over 90% of these were differentially regulated at the whole transcript or exon level across regions and/or time. The majority of these spatiotemporal differences occurred before birth, followed by an increase in the similarity among regional transcriptomes during postnatal lifespan. Genes were organized into functionally distinct co-expression networks, and sex differences were present in gene expression and exon usage. Finally, we demonstrate how these results can be used to profile trajectories of genes associated with neurodevelopmental processes, cell types, neurotransmitter systems, autism, and schizophrenia, as well as to discover associations between SNPs and spatiotemporal gene expression. This study provides a comprehensive, publicly available dataset on the spatiotemporal human brain transcriptome and new insights into the transcriptional foundations of human neurodevelopment.

1.2. Introduction

Human neurodevelopment is a complex and precisely regulated process that unfolds over a protracted period of time[1-3]. Human-specific features of this process are likely to be important factors in the evolution of human specializations[2,3,5,6]. However, in addition to giving us remarkable cognitive and motor abilities, the formation of molecularly distinct and intricate neural circuits may have also increased our susceptibility to certain psychiatric and neurological disorders[4,7-10]. Furthermore, sex differences play an important role in brain development and function and are a risk factor for disorders such as autism spectrum disorders (ASD)[10-14]. Thus, comprehensive knowledge about the spatiotemporal dynamics of the brain transcriptome is essential for a better understanding of neurodevelopment, sexual dimorphism, and evolution, as well as our increased susceptibility to certain brain disorders.

Previous transcriptome studies of the developing human brain have included relatively small sample sizes and predominantly focused on few regions or developmental time points[15-19]. In this study, we explore the transcriptomes of 16 regions from developing and adult postmortem brains of clinically unremarkable donors representing males and females of multiple ethnicities.

1.3. Results

Study design, data generation, and quality control

To investigate the spatiotemporal dynamics of the human brain transcriptome, we created a 15-period system spanning from embryonic development to late adulthood (Table 1). We sampled transient prenatal structures and immature and mature forms of 16 brain regions, including 11 NCX areas, from multiple specimens per period (Table 2; Supplementary Information 2.1; Supplementary Figs. 1-3; Supplementary Table 1). The 11 NCX areas are referred to hereafter collectively as the region NCX. We also genotyped donor's DNA using an Illumina 2.5 million SNP chip (Supplementary Fig. 4; Supplementary Table 2). Only brains from clinically unremarkable donors with no signs of large-scale genomic abnormalities were included in the study (N=57, including 39 with both hemispheres; age, 5.7 post-conceptual weeks to 82 years; sex, 31 males and 26 females; postmortem interval [PMI], 12.11±8.63 [mean±SD] hours; pH, 6.45±0.34 [mean±SD]).

Transcriptome profiling was performed using total RNA extracted from a total of 1,340 dissected tissue samples (RNA integrity number [RIN], 8.83±0.93 [mean±SD]; Supplementary Tables 3 and 4). We used the Affymetrix GeneChip Human Exon 1.0 ST Array platform, which features comprehensive coverage of the human genome, with 1.4 million probe sets that assay expression across the entire transcript or individual exon,

thereby providing redundancy and increasing confidence in estimates of gene-level differential expression (DEX) and differential exon usage (DEU). Descriptions of tissue sampling and quality control measures implemented throughout transcriptome data generation steps are provided in Supplementary Information 2-5 and Supplementary Figures 5-8.

Global transcriptome dynamics

Spatiotemporal gene expression

After quality control metrics were assessed, we performed quantile normalization and summarized core and unique probe sets, representing 17,565 protein-coding genes, into gene-level information. Employing stringent criteria (log_2-transformed signal intensity ≥6 in at least one sample and mean DABG P<0.01 in at least one region and period) to define an "expressed" gene, we found that 15,132 (86.1%) of 17,565 genes surveyed were expressed in at least one brain region during at least one period, and 14,375 (81.8%) were expressed in at least one NCX area (Fig. 1a; Supplementary Information 6.1; Supplementary Fig. 9). To investigate the contribution of different factors to the global transcriptome dynamics, we applied multi-dimensional scaling (MDS) and principal component analysis (PCA), which revealed that regions and age (i.e., spatiotemporal dynamics) contribute more to the global differences in gene expression than other tested variables: sex, ethnicity, or inter-individual variation (Fig. 1b; Supplementary Information 6.2 and 6.3; Supplementary Figs 10 and 11).

To identify genes that were spatially or temporally regulated, we employed a conservative threshold (FDR q-value<0.01 and ≥2-fold log_2-transformed signal intensity difference), included PMI and RIN as technical covariates within all of our ANOVA models of differential expression, considered the influence of dissection variation, and applied a 5-fold jackknife procedure (Supplementary Information 6.4; Supplementary Figs 12 and 13). We found that 70.9% of expressed genes were spatially DEX between any two regions within at least one period, and 24.1% were spatially DEX between any two NCX areas (Fig. 1a). In contrast, 90.0% of expressed genes were temporally DEX between any two periods across regions, and 85.3% were temporally DEX between any two periods across NCX areas. Moreover, 70.0% and 23.9% of expressed genes were both spatially and temporally DEX within brain regions and within NCX areas, respectively. The bulk of spatiotemporal regulation occurred during prenatal development. For instance, 57.7% of NCX-expressed genes were temporally DEX across fetal development (periods 3-7), in contrast to 9.1% during postnatal development (periods 8-12) and a mere 0.7% during adulthood (periods 13-15). Together, these data indicate that the majority of brain-

expressed protein-coding genes are temporally and, to a lesser extent, spatially regulated, and this regulation occurs predominately during development.

Transcriptional architecture of the human brain

To assess the transcriptional relatedness, we calculated correlation matrices of pairwise comparisons between brain regions/NCX areas (Fig. 1c) and performed unsupervised hierarchical clustering across periods 3–15, when all analyzed regions/areas can be consistently followed across time. Among regions, this analysis showed distinct and developmentally regulated clustering of NCX (combination of 11 areas), HIP and AMY, with CBC exhibiting the most distinctive transcriptional profile. At the level of NCX areas, clustering formed the following groups during fetal periods: OFC, DFC, and MFC of the prefrontal cortex; VFC and primary somatomotor cortex (S1C and M1C); and parietal-temporal perisylvian areas (IPC, A1C, and STC). V1C had the most distinctive transcriptional profile of NCX areas throughout development and adulthood. The increased correlations between NCX, HIP, AMY, and the majority of non-V1C NCX areas with age indicate that transcriptional differences are particularly pronounced during development.

Consistent with the clustering observed, CBC showed the greatest number of region-restricted or region-enriched DEX genes with 516 (4.8%) of 10,729 genes spatially DEX (Supplementary Information 6.4; Supplementary Table 5). In contrast, the number of genes highly enriched in the other regions was lower: NCX (46; 0.43%), HIP (48; 0.45%), AMY (4; 0.04%), STR (137; 1.28%), and MD (216; 2.01%). The majority of these spatially enriched genes were also temporally regulated, and some were transiently enriched during a narrow time-window, such as those featured in Supplementary Figures 14 and 15 (NCX: *FLJ32063, KCNS1*; HIP: *CDC20B, METTL7B*; AMY: *TFAP2D, UTS2D;* STR: *C10ORF11, PTPN7*; MD: *CEACAM21, SLC24A5;* CBC: *ESRRB, ZP2)*. These clustering and region-enrichment results reveal that regional transcriptomes are developmentally regulated and likely reflect anatomical differences.

Spatiotemporal differential exon usage

Alternative exon usage is an important mechanism for generating transcript diversity[20,21]. Using the splicing ANOVA and splicing index algorithm with conservative criteria (FDR<0.01 with a minimum 2-fold splice index difference between at least two regions/areas or periods; Supplementary Information 6.5), we found that 13,647 of 15,132 (90.2%) expressed genes exhibited DEU across sampled regions (0.1%), periods (19.5%), or both (70.6%). Of 14,375 NCX genes, 88.7% exhibited DEU across sampled areas (<0.01%), periods (59.8%), or both (28.9%). The regulation of DEU also varied across time, with the vast majority of expressed genes (83.0%) exhibiting temporal DEU across

fetal development, while only 0.9% and 1.4% were temporally regulated across postnatal development and adulthood, respectively.

Focusing on *ANKRD32*, a gene we have previously shown to express an alternative variant in the late midfetal frontal cortex[17], we confirmed and extended DEU findings by revealing that while the longer isoform (*ANKRD32a*) was equally expressed across fetal NCX areas, the shorter isoform (*ANKRD32b*), comprising the last three exons, exhibited dynamic areal patterns. *ANKRD32b* was transiently expressed in a gradient along the anterior-posterior axis of the midfetal frontal cortex, with the highest expression in OFC and lowest in M1C. Prior to this, *ANKRD32b* was most highly enriched in the ITC and, to a lesser extent, in the STC. These spatiotemporal patterns disappeared after birth, when only *ANKRD32a* was expressed, and were not observed in the mouse NCX of equivalent periods (Supplementary Fig. 16; Supplementary Table 6). These findings illustrate the complexity of DEU in the human brain and demonstrate how specific alternative transcripts can be spatially restricted during a narrow developmental window and with inter-species differences.

Sex differences in the transcriptome

Sex-biased gene expression

Previous studies have identified sexually dimorphic gene expression in the developing and adult human brain[12-14]. Analysis of our dataset using a sliding window algorithm and t-test model (FDR<0.01 with >2-fold difference in $\log_2$-transformed signal intensity; Supplementary Information 6.6), identified 159 genes, including a number of previously reported and newly uncovered genes with male- or female-bias in expression located on the Y (13), X (9), and autosomal (137) chromosomes. A large fraction (76.7%) displayed male-biased expression (Fig. 2a; Supplementary Table 7). Notable spatial differences were observed, and more genes had sex-biased expression during prenatal development than during postnatal life, with the adult brain characterized by the lowest number.

Consistent with previous findings[13,14], we found that the largest differences were attributable to Y chromosome genes, especially *PCDH11Y*, *RPS4Y1*, *USP9Y*, *DDX3Y*, *NLGN4Y*, *UTY*, *EIF1AY*, and *ZFY*, which displayed constant expression across regions and periods, with the exception of *PCDH11Y* down-regulation in the postnatal CBC (Fig. 2b). Interestingly, the functional homologues of these genes on the X chromosome (*PCDH11X*, *RPS4X*, *USP9X*, *DDX3X*, *NLGN4X*, *UTX*, and *EIF1AX*), except *ZFX* during fetal development, were not up-regulated in a compensatory manner in female brains (Supplementary Fig. 17).

We also found other X-linked and autosomal genes with sex-biased expression and distinct spatiotemporal patterns, including functionally uncharacterized transcripts

(*LOC554203, C3ORF62, FLJ35409*, and *DKFZP586*), *S100A10* (which has been linked to depression[22]), and *IGF2* (an imprinted autosomal gene previously implicated in brain growth and cognitive function[23,24]) that exhibited population-level male-biased developmental expression (Fig. 2c).

Sex-biased exon usage

We next explored sex-biased DEU using a sliding window algorithm and splicing t-test model (FDR<0.01 and splicing index>2; Supplementary Information 6.6). We identified 155 genes (145 autosomal) that displayed sex-biased expression of probe sets encoding one or a subset of exons (Supplementary Table 8) in one or multiple regions/NCX areas. These included several members of the collagen family of genes (*COL1A1, COL1A2, COL3A1, COL5A2,* and *COL6A3*), *C3, KCNH2* (a gene associated with schizophrenia[25]), *NOTCH3* (a gene mutated in a common form of hereditary stroke disorder[26]), *ELN* (a gene located within the Williams syndrome critical region[27]), and *NLGN4X* (a X chromosome gene implicated in synapse function and associated with ASD and moderate X-linked intellectual disability[11,28]). Although comparably expressed in males and females at the population and gene-level (Supplementary Fig. 17), *NLGN4X* exhibited a significant male-bias in expression of exon 7 and, to a lesser extent, exons 1, 5 and 6 in a developmentally regulated manner (Fig. 3). Together, these findings show that developmentally and spatially regulated differences in gene- and exon-level expression exist between male and female brains.

Co-regulated transcriptional networks

To extract additional biological information embedded in the multi-dimensional transcriptome dataset, we performed weighted gene co-expression network analysis[29] to identify modules of co-regulated genes. We identified 29 modules associated with distinct spatiotemporal expression patterns and biological processes (Fig. 4a; Supplementary Information 6.7; Supplementary Table 9-11; Supplementary Figs 18-20). Among modules corresponding to specific spatiotemporal patterns, M8 consisted of 24 genes with a common developmental trend that exhibited highest levels in early fetal NCX and HIP (period 3) and then progressively declined with age until infancy (period 9) (Fig. 4b). In contrast, M15 contained 310 genes exhibiting changes in the opposite direction in the NCX, HIP, AMY, and STR (Fig. 4c). Gene ontology (GO) enrichment analysis showed that M8 genes were enriched for GO categories related to neuronal differentiation (Bonferroni-adjusted $P=7.7\times10^{-3}$) and transcription factors ($P=5.2\times10^{-3}$) (Supplementary Information 6.8; Supplementary Table 9).

Genes with the highest degree of connectivity within a module are termed hub genes and are expected to be functionally important within the module. M8 hub genes included transcription factors *TBR1, FEZF2, FOXG1, SATB2, NEUROD6,* and *EMX1* (Fig. 4b), which have been functionally implicated in NCX and HIP development[4,30-38]. Furthermore, *FOXG1* variants have been linked to Rett syndrome and intellectual disability[34]. Conversely, M15 GO categories included ionic channels (P=8.0×10^{-8}) and neuroactive ligand-receptor interaction (P=4.0×10^{-14}). Sequence variants in hub genes (Fig. 4c) have been linked to major depression (*GDA*)[39] and to schizophrenia and affective disorders (*NRGN* and *RGS4*)[7,40].

We also identified two large-scale temporally regulated modules (M20 and M2) with opposite developmental trajectories of genes co-expressed across regions, which either gradually declined (M20) or increased (M2) in expression with age (Supplementary Figs 21 and 22). M20 was enriched for GO categories related to zinc finger proteins (P=7.3×10^{-48}) and transcription factors (P=4.8×10^{-50}), including many ZNF and SOX family members, while M2 was enriched for categories related to membrane proteins (P=1.8×10^{-21}) and genes involved in calcium signaling (P=8.1×10^{-10}), synaptic transmission (P=1.6×10^{-6}), and neuroactive ligand-receptor interaction (P=4.1×10^{-4}), reflecting processes important in postnatal brain maturation. Their hub genes encoded transcriptional factors, modulators of chromatin state, and signal transduction proteins, all of which likely play an important role in driving the networks. Dramatic expression shifts of M20 and M2 in the opposite direction just before birth indicate that this period is associated with transcriptional changes that likely reflect environmental influences on brain development and intrinsic changes in cellular composition and functional processes.

Expression trajectories of neurodevelopment

One important application for the generated dataset is to gain insight into normal and abnormal human neurodevelopment by analyzing trajectories of individual genes or groups of genes associated with a particular neurobiological category or disease. To test this strategy, we compared our expression data of *DCX* (a gene expressed in neuronal progenitor cells and migrating immature neurons), as well as genes associated with dendrite (*MAP1A, MAPT, CAMK2A*) and synapse (*SYP, SYPL1, SYPL2, SYN1*) development, with independently generated, non-transcriptome human datasets. The *DCX* expression trajectory was remarkably reminiscent of the reported changes in the density of DCX-immunopositive cells in the postnatal human hippocampus (Fig. 5a)[36,40]. In our transcriptome dataset, *DCX* expression progressively increased until early midfetal development (period 5) and then gradually declined with age until early childhood (period 10). Likewise, expression trajectories of dendrite and synapse development gene groups

closely paralleled the growth of basal dendrites of DFC pyramidal neurons[41] (Spearman correlation, r=0.810 for layer 3 and r=0.700 for layer 5; Fig. 5b) and DFC synaptogenesis[42] (r=0.946; Fig. 5c), respectively. Steep increases in both processes occurred between late midfetal and late infancy, indicating that a considerable portion of these two processes occurs before birth and reaches a plateau around late infancy.

After demonstrating the accuracy and viability of using the dataset to profile human neurodevelopment, we manually curated lists of genes associated with over 80 categories, including various neurodevelopmental processes, neural cell types, and neurotransmitter systems (Supplementary Information 6.9; Supplementary Table 12). Interesting trajectories and notable differences in their onset, rates of incline and decline, and shapes were observed within and between brain regions for in the categories, including major neurodevelopmental processes (neural cell proliferation and migration, dendrite and synapse development, and myelination; Fig. 5d), cortical GABAergic inhibitory interneurons (*CALB1, CALB2, NOS1, PVALB,* and *VIP*), and glutamate receptors (Supplementary Figs 23 and 24). As expected, two major trends were observed in neurodevelopmental trajectories: changes in expression of cell proliferation genes preceded the increase in expression of *DCX*, both of which decreased during perinatal development while synapse development, dendrite development, and myelination trajectories increased. Interestingly, the NCX expression trajectory for synapse development (Fig. 5c, d) did not display a drastic decline during late childhood or adolescence as previously reported for synapse density[42]. We also identified network modules and additional genes that are highly correlated with the categories (Supplementary Tables 13 and 14). For example, M20 and M2 were strongly correlated with neuron migration (r=0.894) and myelination (r=0.972), respectively.

In addition, our dataset allowed us to generate expression trajectories of genes commonly associated with ASD and schizophrenia. We investigated a number genes previously linked to these disorders (Supplementary Information 6.10) and observed distinct and dynamic expression patterns, especially among NCX areas, including examples for *CNTNAP2*, *MET*, *NLGN4X*, and *NRGN* shown in Supplementary Figure 25. To gain insight into potential biological functions of ASD and schizophrenia-associated genes in human neurodevelopment, we identified other genes with significantly correlated spatiotemporal expression profiles and performed GO enrichment analysis (Supplementary Tables 15 and 16). These findings reveal spatiotemporal differences associated in these expression trajectories and provide additional co-expressed genes that can be interrogated for their role in the respective processes or disorders.

Expression quantitative trait loci

Previous reports have demonstrated the connection between SNP-based genetic variation and expression quantitative trait loci (eQTLs) in the adult postmortem human brain, primarily in the cerebral cortex[43-47]. Our multiregional developmental dataset enabled us to search for association between SNP genotypes and spatiotemporal gene expression. We tested only for *cis*-eQTLs, restricting the search to SNPs within 10 kb of either transcription start site (TSS) or termination site (TES), as opposed to *trans*-eQTLs, which would require much larger sample sizes.

Implementing a conservative strategy (gene-wide Bonferroni correction followed by genome-wide FDR<0.1; Supplementary Information 9), we identified 39 NCX, 8 HIP, 4 AMY, 2 STR, 6 MD, and 5 CBC genes (Supplementary Table 17) with evidence of *cis*-eQTL, including two previously reported genes (*ITGB3BP* and *ANKRD27*)[45,47]. Consistent with previous studies[48], associated SNPs were enriched near TSS or TES (Fig. 6a).

An example of a significant association in NCX, MD, and CBC is rs10785190 and *GLIPR1L2*, a member of the glioma pathogenesis-related 1 family of genes[49]. The expression differences were observed at the level of the whole transcript and exons 1 and 2, the only exons we observed to be expressed at appreciable levels in the NCX (Fig. 6b). The NCX likely had more *cis*-eQTLs than other regions due to its smaller variation in gene expression resulting from the averaged expression of 11 areas. Many eQTLs identified as significant in NCX also have similar associations in other regions, though they were not statistically significant after the conservative genome-wide correction (Supplementary Table 17). Thus, we have identified polymorphic regulators of transcription in different regions across development, potentially providing insights into inter-individual differences and genetic control of the brain transcriptome.

1.4. Conclusions

We have described the generation and analysis of the transcriptome and corresponding genotyping dataset covering multiple regions of the developing and adult brain. This dataset (available at www.humanbraintranscriptome.org), along with another recent study[50], provides opportunities for a variety of further investigations and comparisons with other transcriptome-related datasets of both healthy and diseased states.

Our analysis revealed several important aspects of the human brain transcriptome and expands current knowledge of the transcriptional events in human neurodevelopment. Consistent with many previous studies in other species, we found that gene expression and exon usage have complex, dynamically regulated patterns across time and space and are far more prominent than differences between sexes, ethnicities, or individuals despite their underlying genomic differences.

We confirmed and expanded on previous findings of sexual dimorphism in gene expression and exon usage, including several disease-related genes. These findings offer possible transcriptional mechanisms underlying sex differences in the incidence, prevalence, and severity of many disorders. We also demonstrated how the dataset can be used to profile trajectories of genes associated with specific neurobiological categories or disorders, many of which would not likely be evident in the transcriptomes of commonly studied model systems. Coupled with analysis of co-expressed genes in the dataset, these provide information about when and where these genes are expressed in the brain, which can help infer their function. Our data can enhance genome-wide association and linkage studies by narrowing the focus to the candidate genes that are specifically expressed during development or restricted to a specific region known to be preferentially affected.

We show associations between specific SNPs and expression levels in different regions of the developing human brain, indicating that genetic variation contributes to inter-individual transcriptome variability across regions and development. While the current number of specimens in our dataset restricted our power to detect many of the possible eQTLs, we have identified a set of *cis*-eQTLs, including many not previously reported, that may provide insights into expression-regulatory elements operating in the brain.

Although these findings highlight the complexity of gene expression and exon usage in the human brain, there are several potential limitations in our data that warrant discussion. Foremost, we used stringent criteria in order to faithfully characterize general transcriptional patterns and minimize false positives, rather than to capture all the changes that may occur. Also, we analyzed dissected tissue samples that contain multiple cell types, thus diluting the genetic contribution and dynamic range of genes expressed in a more cell type specific manner. Current limitations prevent us from using cell type specific approaches in systematically analyzing the spatiotemporal transcriptome. Furthermore, the number of brains and regions analyzed so far is not sufficient to fully investigate the whole magnitude of transcriptional changes or the full range of eQTLs. Application of sequencing technology will allow even more in-depth analysis of the transcriptome and discovery of novel and low-expressing transcripts and their associated spatiotemporal patterns. Finally, while specific patterns of expression are often linked to equally specialized biological processes, it is important to keep in mind that the relationship between mRNA and protein levels is not always linear or translated into apparent phenotypic differences. As these concerns are addressed in the future with the addition of samples to our dataset and the generation of new datasets from human and nonhuman brains throughout development, it will be possible to uncover further insights into the transcriptional foundations of human brain development and evolution.

1.5. Methods summary

The Supplementary Information provides a full description of materials and methods, including tissue acquisition and processing, data generation, validation, and analyses.

1.6. Acknowledgements

We thank A. Belanger, V. Imamovic, R. Johnson, P. Larton, S. Lindsay, B. Poulos, J. Rajan, D. Rimm, and R. Zielke for assistance with tissue acquisition, D. Singh for technical assistance, I. Kostovic and Z. Petanjek for Golgi measurements, P. Levitt for suggesting the ITC, and D. Karolchik and A. Zweig for help in creating tracks for the UCSC Genome browser. We also thank A. Beckel-Mitchener, M. Freund, M. Gerstein, D. Geschwind, T. Insel, M. Judas, J. Knowles, E. Lein, P. Levitt, N. Parikshak and members of the Sestan laboratory for discussion and criticism. Tissue was obtained from several sources including the Human Fetal Tissue Repository at the Albert Einstein College of Medicine, the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, the Laboratory of Developmental Biology at the University of Washington (supported by HD000836 grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development), and the Joint MRC/Wellcome Trust Human Developmental Biology Resource (http://hdbr.org) at the IHG, Newcastle Upon Tyne (UK funding awards G0700089 and GR082557). Support for pre-doctoral fellowships was provided by the China Scholarship Council (Y.Z.), the Portuguese Foundation for Science and Technology (A.M.M.S.), the Samsung Scholarship Foundation (Y.S.), Fellowship of the German Academic Exchange Service – DAAD (S.M.), and the NIDA grant DA026119 (T.G.). This work was supported by grants from the US National Institutes of Health (MH081896, MH089929, NS051869), Kavli Foundation, NARSAD, and the James S. McDonnell Foundation Scholar Award (N.S.).

1.7. Tables

Table 1 | Periods of human development and adulthood as defined in this study

Period	Description	Age
1	Embryonic	4≤ Age <8 Postconceptual weeks (PCW)
2	Early fetal	8≤ Age <10 PCW
3	Early fetal	10≤ Age <13 PCW
4	Early midfetal	13≤ Age <16 PCW
5	Early midfetal	16≤ Age <19 PCW
6	Late midfetal	19≤ Age <24 PCW
7	Late fetal	24≤ Age <38 PCW
8	Neonatal and early infancy	Birth≤ Age <6 Postnatal months (M)
9	Late infancy	6 M≤ Age <12 M
10	Early childhood	1≤ Age <6 Postnatal years (Y)
11	Middle and late childhood	6≤ Age <12 Y
12	Adolescence	12≤ Age <20 Y
13	Young adulthood	20≤ Age <40 Y
14	Middle adulthood	40≤ Age <60 Y
15	Late adulthood	60 Y ≤ Age

Table 2 | Ontology and nomenclature of analyzed brain regions and NCX areas

Periods 1-2	Periods 3 - 15
FC: Frontal cerebral wall	**OFC**: Orbital prefrontal cortex
	DFC: Dorsolateral prefrontal cortex
	VFC: Ventrolateral prefrontal cortex
	MFC: Medial prefrontal cortex
	M1C: Primary motor (M1) cortex
PC: Parietal cerebral wall	**S1C**: Primary somatosensory (S1) cortex
	IPC: Posterior inferior parietal cortex
TC: Temporal cerebral wall	**A1C**: Primary auditory (A1) cortex
	STC: Superior temporal cortex
	ITC: Inferior temporal cortex
OC: Occipital cerebral wall	**V1C**: Primary visual (V1) cortex
HIP: Hippocampal anlage	**HIP**: Hippocampus
	AMY: Amygdala
VF: Ventral forebrain	**STR**: Striatum
MGE: Medial ganglionic eminence	
LGE: Lateral ganglionic eminence	
CGE: Caudal ganglionic eminence	
DIE: Diencephalon	**MD**: Mediodorsal nucleus of the thalamus
DTH: Dorsal thalamus	
URL: Upper (rostral) rhombic lip	**CBC**: Cerebellar cortex

1.8. Figures

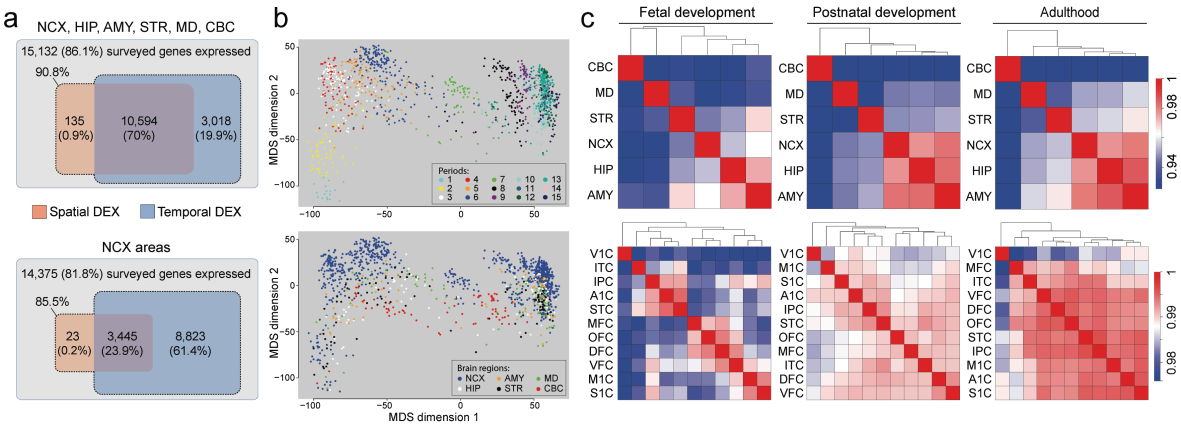


Figure 1 | Global spatiotemporal dynamics of gene expression. a, Venn diagrams representing total number of genes considered to expressed and the number of spatially and temporally DEX genes for brain regions (top) and NCX areas (bottom). **b**, MDS plot showing transcriptional similarity, colored by periods (top) and regions (bottom). Euclidean distance of $\log_2$-transformed signal intensity (nonmetric, stress=18.9%) was used to measure pairwise similarity. **c**, Heat map matrix of pairwise Spearman correlations between brain regions (top) or NCX areas (bottom) during fetal development (periods 3–7), postnatal development (periods 8–12), and adulthood (periods 13–15).

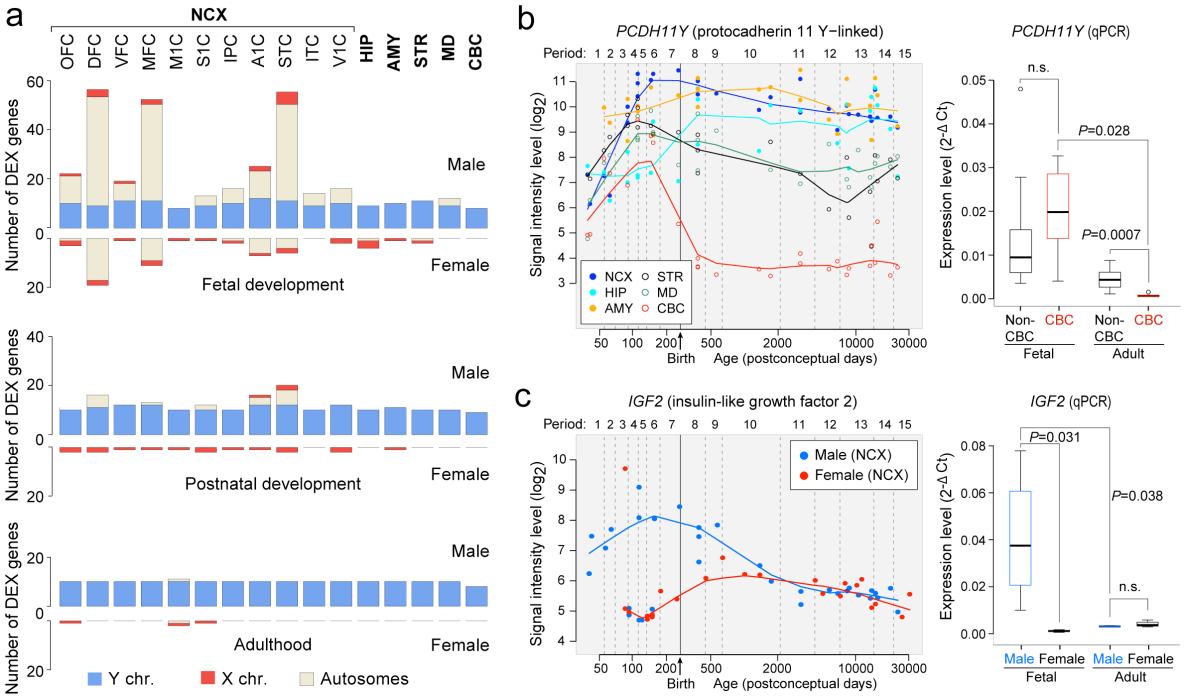


Figure 2 | Sex-biased gene expression. **a,** Number of sex-biased DEX genes in brain regions/NCX areas during fetal development (periods 3–7), postnatal development (periods 8–12), and adulthood (periods 13–15). **b,** *PCDH11Y* exon array signal intensity (left) and qRT-PCR validation (right) (N=5 male brains per period). **c,** *IGF2* exon array signal intensity (left) and qRT-PCR (right) validation in NCX (N=4 per sex and period). *P*-values were calculated by unpaired t-test. Whiskers indicate 5th and 95th percentile.

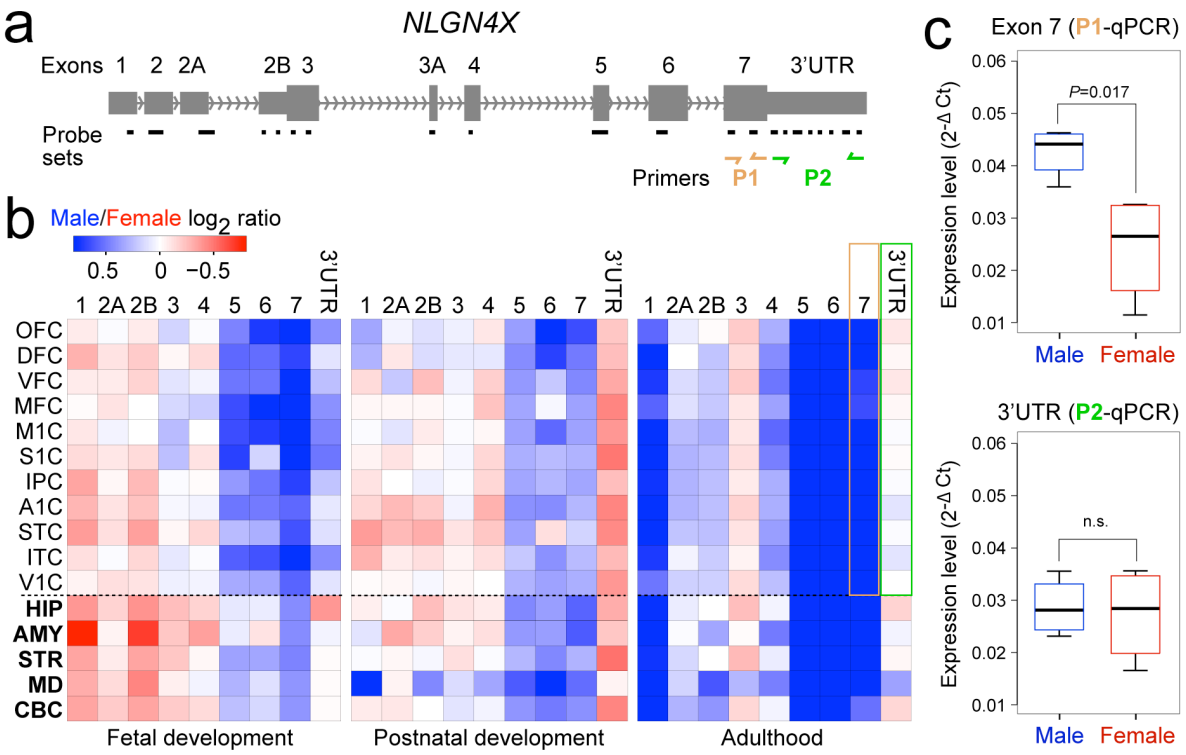


Figure 3 | Sex-biased differential exon usage. **a,** Gene structure and probe set composition of *NLGN4X*. Depicted by yellow and green arrows are primers used for qRT-PCR validation. **b,** Heat map of the $\log_2$ male/female signal intensity ratio of each exon for fetal development (periods 3–7), postnatal development (periods 8–12), and adulthood (periods 13–15). Differences in expression of exon 7 (yellow frame) and 3'UTR (green frame) in adult NCX are highlighted. Note that exons 2 and 3A did not meet our expression criteria and are not represented. **c,** qRT-PCR validation of exon 7 and 3'UTR expression in adult NCX (N=4 per sex). *P*-values were calculated by unpaired t-test. Whiskers indicate 5th and 95th percentile.

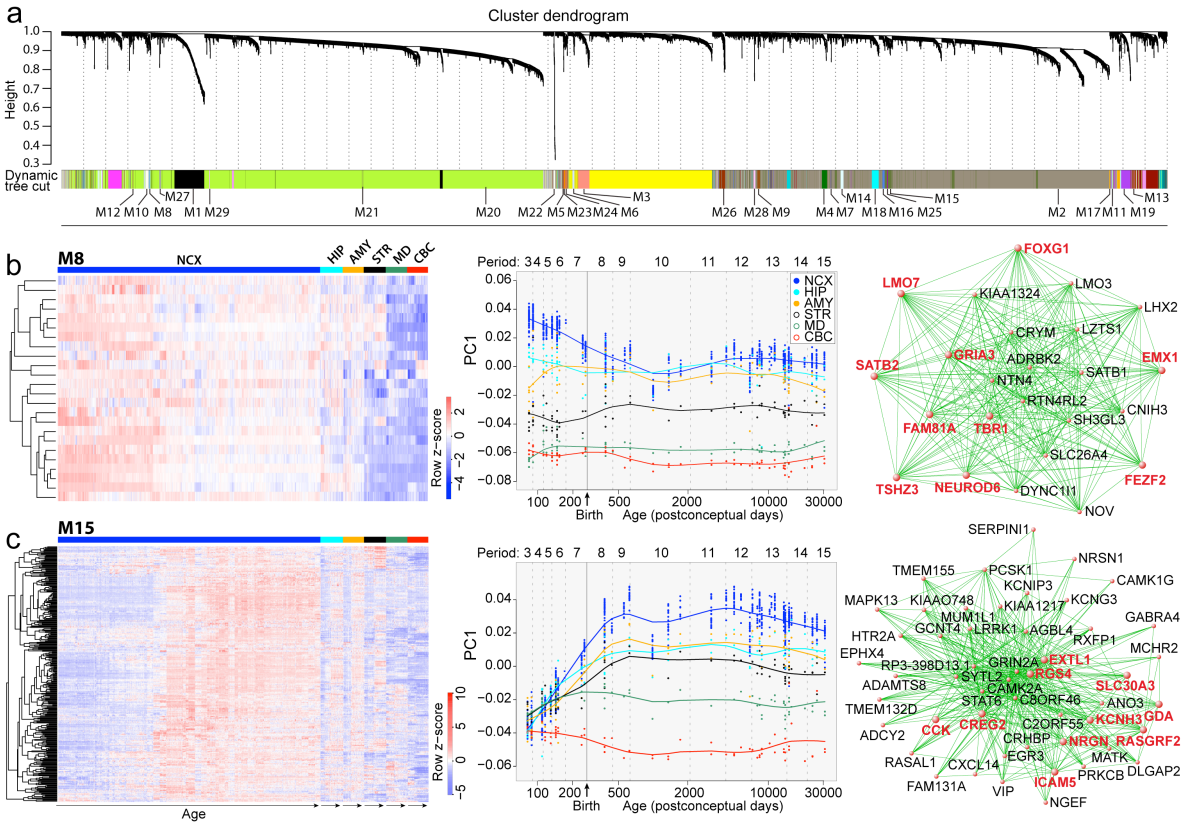


Figure 4 | Global co-expression networks and gene modules. a, Dendrogram from gene co-expression network analysis of samples from period 3 to 15. Modules of co-expressed genes were assigned a color and number (M1 to M29). **b**, Heat map of genes in M8 (left) showing the spatiotemporal expression pattern after hierarchical clustering. The expression values for each gene are arranged in the heat map, ordered first by brain regions, then by age and last by NCX areas. Spatiotemporal pattern of M8 (middle) summarized by the first principal component (PC1) for expression of genes in the module across age. Top 50 M8 genes (right), defined by the highest intramodular connectivity, Top 10 hub genes are in red. **c,** Same analyses were performed for M15. Results for other modules are available in Supplemental Information.

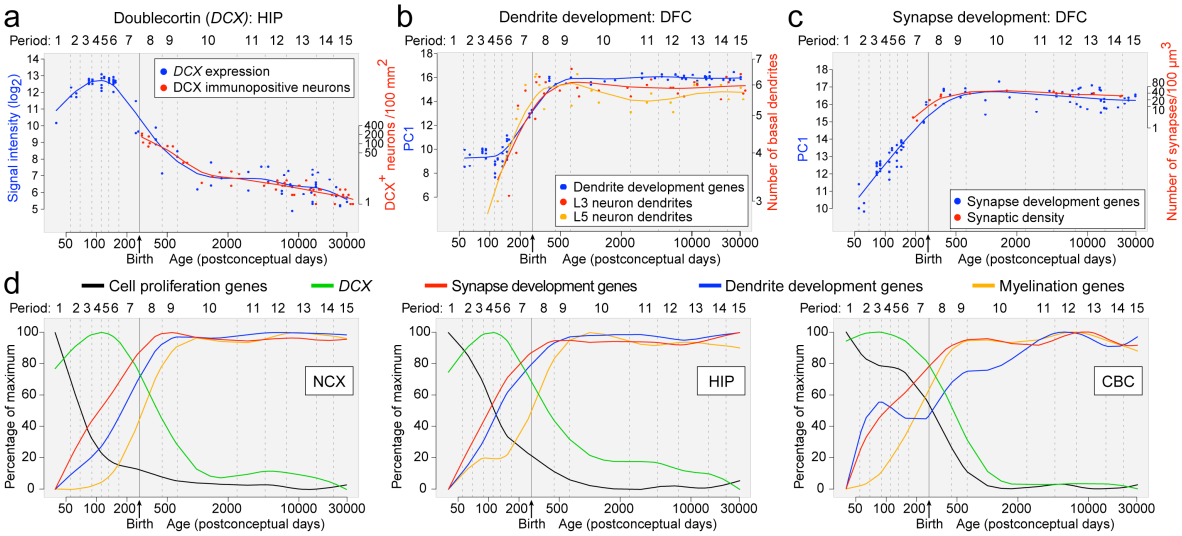


Figure 5 | Trajectories of genes associated with neurodevelopmental processes. **a**, Comparison between *DCX* expression in HIP and the density of DCX-immunopositive cells in the human dentate gyrus[36]. **b**, Comparison between transcriptome-based dendrite development trajectory in DFC and Golgi method-based growth of basal dendrites of layer 3 (L3) and 5 (L5) pyramidal neurons in the human DFC[41]. **c**, Comparison between transcriptome-based synapse development trajectory in DFC and density of DFC synapses calculated using electron microscopy[42]. For (b) and (c), PC1 for gene expression was plotted against age to represent the developmental trajectory of genes associated with dendrite or synapse development. Independent datasets were centered, scaled, and plotted on a logarithmic scale. **d,** PC1 value percentage of maximum PC1 value for given set of genes plotted against age to represent general trends and regional differences in several neurodevelopmental processes in NCX, HIP, and CBC.

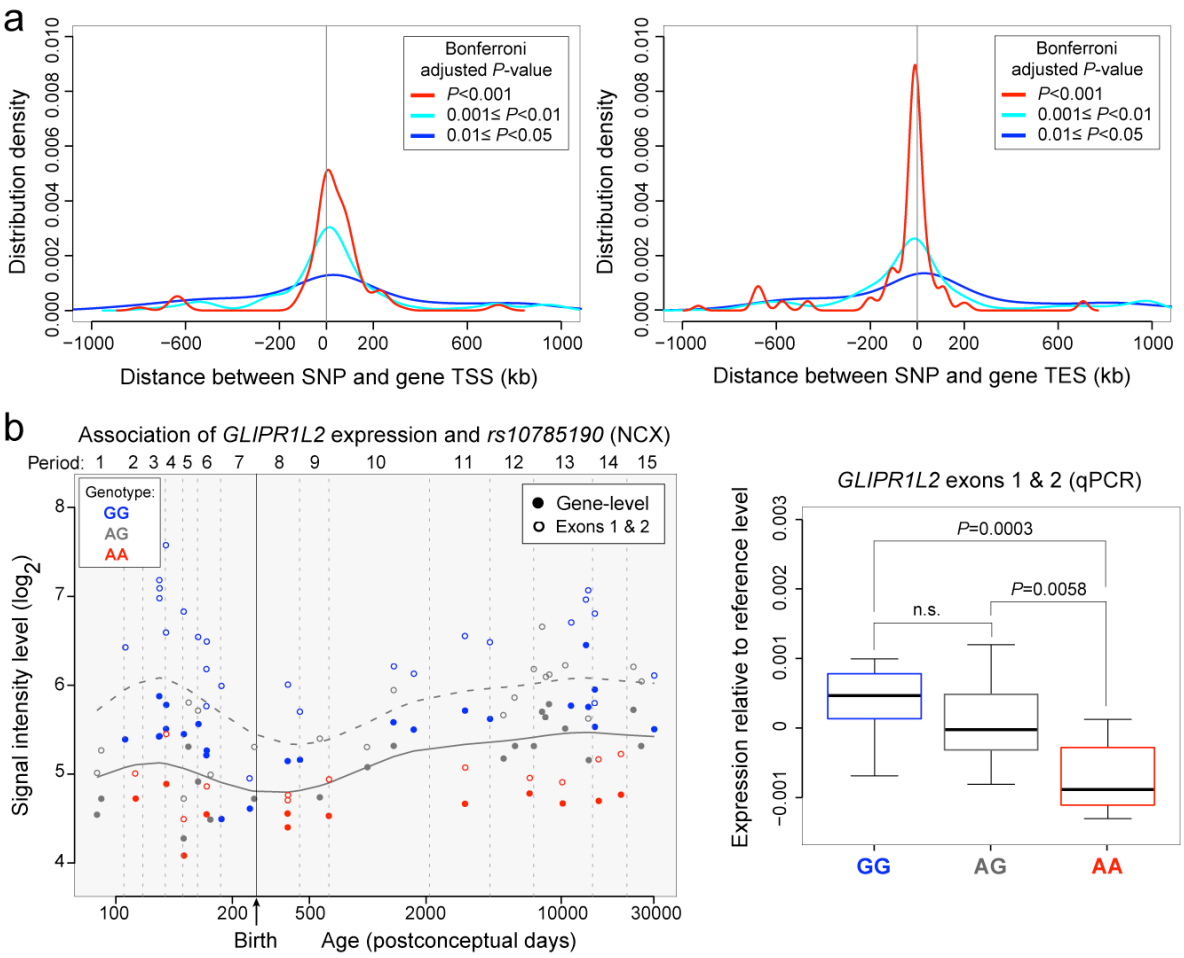


Figure 6 | Association between SNPs and gene expression. a, **b**, SNP distribution around TSS (a) and TES (b) of the associated genes based on several Wald test *P*-value cut offs after gene-wide Bonferroni correction. **c**, *GLIPR1L2* expression association with rs10785190 genotype, a SNP located in exon 1. The horizontal solid and dashed LOWESS curves are the developmental trend of gene and exon 1 and 2 expression, respectively. **d**, qRT-PCR validation of exon 1 and 2 expression in NCX for each genotype (N=14 GG, 14 AG, 8 AA) plotted relative to the local regression smooth curve to facilitate comparison across developmental periods. *P*-values were calculated by unpaired t-test. Whiskers indicate 5th and 95th percentile.

1.9. Supplementary Information

Table of Contents

1. Introduction

2. Supplementary Information on the Study Design

- 2.1. Definition of periods of human development and adulthood used in this study
- 2.2. Ontology and anatomical definition of sampled brain regions and NCX areas
  - 2.2.1. Neocortex (NCX)
    - 2.2.1.1. Frontal cortex (FC): OFC, DFC, VFC, MFC, M1C
    - 2.2.1.2. Parietal cortex (PC): S1C, IPC
    - 2.2.1.3. Temporal cortex (TC): A1C, STC, ITC
    - 2.2.1.4. Occipital cortex (OC): V1C
  - 2.2.2. Hippocampus (HIP)
  - 2.2.3. Amygdala (AMY)
  - 2.2.4. Ventral forebrain (VF)
    - 2.2.4.1. Medial ganglionic eminence (MGE)
    - 2.2.4.2. Lateral ganglionic eminence (LGE)
    - 2.2.4.3. Caudal ganglionic eminence (CGE)
    - 2.2.4.4. Striatum (STR)
  - 2.2.5. Diencephalon (DIE)
    - 2.2.5.1. Dorsal thalamus (DTH)
    - 2.2.5.2. Mediodorsal nucleus of the thalamus (MD)
  - 2.2.6. Upper (rostral) rhombic lip (URL)
    - 2.2.6.1. Cerebellar cortex (CBC)
- 2.3. Workflow and quality control measures

3. Tissue Procurement and Sampling

- 3.1. Tissue procurement
- 3.2. Neuropathological evaluation
- 3.3. Selection criteria for brain specimens
- 3.4. Tissue processing and dissections
  - 3.4.1. Tissue dissection methods
    - 3.4.1.1. Regional sampling from fresh brain specimens
    - 3.4.1.2. Regional sampling from frozen brain specimens
    - 3.4.1.3. Regional sampling from specimens processed in RNA*later*

3.4.2. Histological verification of tissue sampling
3.4.3. Dissection scoring
3.4.4. Tissue pulverization

4. DNA Isolation and Genotyping Data Analyses

4.1. DNA extraction and genotyping
4.2. Copy number variation (CNV) analysis and genomic quality control
4.3. Corroborating and refining ethnic background from genotypes

5. RNA Isolation, Exon Array Processing, and Quality Assesment

5.1. RNA extraction
5.2. Exon array hybridization
5.3. Exon array quality assessment methods
5.4. Detection of outliers
5.5. Exon array data pre-processing

6. Exon Array Data Analyses

6.1. Gene-level analysis
6.2. Multidimensional scaling (MDS)
6.3. Principal component analysis (PCA)
6.4. Identification of spatial and temporal DEX genes
6.5. Analysis of DEU
6.6. Identification of sex-bias in DEX and DEU
6.7. Weighted gene co-expression network analysis
6.7.1. Dataset filtering
6.7.2. Network construction and module detection
6.7.3. Module filtering
6.7.4. Comparison of modules with transcriptional profiling of neurobiological categories
6.8. Gene ontology (GO) enrichment analysis
6.9. Creation of gene lists for and transcriptional profiling of neurobiological categories
6.10. Expression trajectories and gene correlations for ASD- and schizophrenia-associated genes

7. Transcriptome Validation Methods

7.1. Quantitative real-time RT-PCR and semi-quantitative RT-PCR
7.2. List of PCR primers used in this study

7.3. Immunohistochemistry and histochemistry

8. Dissection of the Mouse Neocortex

9. Expression Quantitative Trait Loci Analysis

10. Supplementary Tables

Supplementary Table 1 | Donor/specimen metadata

Supplementary Table 2 | List of CNVs per specimen

Supplementary Table 3 | Tissue sample metadata

Supplementary Table 4 | List of exon array CEL files

Supplementary Table 5 | List of region-enriched DEX genes

Supplementary Table 6 | Tukey's pairwise analysis of human and mouse *ANKRD32* isoforms

Supplementary Table 7 | List of sex-biased DEX genes

Supplementary Table 8 | List of sex-biased DEU genes

Supplementary Table 9 | Co-expression network modules

Supplementary Table 10 | Correlation between network modules and neurobiological categories

Supplementary Table 11 | Correlation between network modules and confounders

Supplementary Table 12 | Gene lists for neurobiological categories

Supplementary Table 13 | Genes correlated with genes involved in neurodevelopmental processes

Supplementary Table 14 | Genes correlated with genes expressed in specific neural cell types

Supplementary Table 15 | Transcripts correlated with ASD-associated genes

Supplementary Table 16 | Transcripts correlated with schizophrenia-associated genes

Supplementary Table 17 | Regional *cis*-eQTLs

11. Supplementary Figures

Supplementary Fig. 1 | Demarcation of adult brain regions and NCX areas

Supplementary Fig. 2 | Demarcation of fetal brain regions and NCX areas

Supplementary Fig. 3 | Histological evaluation and dissection of NCX areas

Supplementary Fig. 4 | Demographics, genotyping, and RNA integrity of analyzed brain specimens

Supplementary Fig. 5 | Quality control workflow

Supplementary Fig. 6 | Exon array hybridization quality control

Supplementary Fig. 7 | Expression analysis with and without SNP-containing probe sets

Supplementary Fig. 8 | Hierarchical clustering of NCX samples

Supplementary Fig. 9 | Distribution plot of expressed genes

Supplementary Fig. 10 | Multidimensional scaling according to sex, ethnicity, and donor

Supplementary Fig. 11 | Principal component analysis

Supplementary Fig. 12 | Correlations between PMI, pH, RIN, DS, and gene expression

Supplementary Fig. 13 | Five-fold jackknife procedure

Supplementary Fig. 14 | Prenatal region-enriched genes

Supplementary Fig. 15 | Postnatal region-enriched genes

Supplementary Fig. 16 | Spatiotemporal analysis of alternative exon usage of *ANKRD32*

Supplementary Fig. 17 | Expression of Y and X chromosome homologues in male and female brain

Supplementary Fig. 18 | WGCNA Module 9

Supplementary Fig. 19 | WGCNA Module 19

Supplementary Fig. 20 | WGCNA Module 23

Supplementary Fig. 21 | WGCNA Module 20

Supplementary Fig. 22 | WGCNA Module 2

Supplementary Fig. 23 | Expression trajectories of genes associated with specific cortical GABAergic interneuron subclasses

Supplementary Fig. 24 | Expression trajectories of genes encoding the subunits of glutamate receptors

Supplementary Fig. 25 | Spatiotemporal expression patterns of disease-related genes

1. Introduction

In this Supplementary Information we provide further information regarding the study design, materials and methods, and additional data. The materials and methods section provide detailed description of the collection, dissection methods, and quality control assessments of postmortem human brain tissue used in this study. We provide technical descriptions of data generation and analyses using different platforms. We also make available additional data that were discussed in the main manuscript. Finally, we present supplementary figures and tables generated from sample metadata and specific gene lists.

2. Supplementary Information on the Study Design

2.1. Definition of periods of human development and adulthood used in this study

We created a classification system of human brain development and adulthood comprised of 15 time periods, particularly emphasizing the timing of specific neurodevelopmental processes in the cerebral cortex, a structure that is central to the most distinctive traits of human cognition. To avoid confusion with the well-known Carnegie stages for human embryonic development[51], we divided our specimens into developmental periods covering both prenatal and postnatal development and adulthood. Our system took into consideration previous developmental classifications of the human brain[51-56].

2.2. Ontology and anatomical definition of sampled brain regions and NCX areas

Brain development is a highly dynamic process during which each region undergoes distinct organizational and maturational changes. Thus we created a structural ontology that contains brain structures (e.g., NCX areas, HIP, AMY, STR, MD, CBC) that are well defined throughout most of time periods and several transient structures (e.g., MGE, LGE, CGE, URL). In total, 8 structures were analyzed for period 1, 10 regions for period 2 and up to 16 regions for periods 3–15. Below we describe this ontology and anatomical definition of sampled brain regions and NCX areas based on histological verification.

2.2.1. Neocortex (NCX)

Samples collected from period 1 and 2 specimens contained the entire thickness of the cerebral wall. Samples collected from period 3–7 specimens contained the marginal zone, cortical plate, and part of the underlying subplate (Supplementary Fig. 3). Samples from period 8–15 specimens were dissected such that the entire gray matter (layer 1-6) and part of the underlying subplate (periods 8 and 9) or white matter (periods 10 – 15) were collected (Supplementary Fig. 3). Nissl staining of the neighbouring thin block was used to histologically verify the identity of the dissected area and to microscopically evaluate

tissue. Neocortical cytoarchitecture of each sample was compared to areal cytoarchitectonic maps to distinguish Brodmann areas (BA)[57]. Samples with incorrect cytoarchitecture or abnormal microscopical appearance were excluded from the study. Neocortical areas (see below) were grouped according to the lobes from which they were sampled.

2.2.1.1. Frontal cortex (FC)

For ***period 1***, the sampled area corresponded to the anterior third of telencephalic vesicle (cerebral wall) corresponding to prospective FC.

For ***period 2***, the sampled area corresponded to different parts (orbital (OFC), dorsolateral (DFC), ventrolateral (VFC), and medial (MFC) of the anterior part of telencephalic vesicle (cerebral wall) corresponding to prospective FC. In addition, paracentral region corresponding approximately to the prospective motor and parietal somatosensory (M1C/S1C) cerebral wall was dissected as one sample (MSC).

For ***periods 3–7***, prior to the appearance of all gyri and sulci, multiple areas of the FC were sampled as follows (Supplementary Fig. 2):

- **Orbital prefrontal cortex (OFC)** was sampled from the middle part of the orbital surface of the cerebral hemisphere, immediately next to the prospective gyrus rectus.
- **Dorsolateral prefrontal cortex (DFC)** was sampled from the middle third of the dorsolateral surface of the anterior third of the cerebral hemisphere.
- **Ventrolateral prefrontal cortex (VFC)** was sampled from the posterior part of the frontal operculum, above the lateral sulcus and prospective insula.
- **Medial prefrontal cortex (MFC)** was sampled from the perigenual and subgenual region of the medial surface.
- **Primary motor cortex (M1C)**, prior to the appearance of the central sulcus, was sampled from the anterior third of the middle third of the cerebral hemisphere, medial third and upper part of the lower third of the dorsolateral surface. We used the striatum at the septal level as the landmark between the anterior and middle one third of the dorsolateral cortical surface. In some cases, we sampled M1C and S1C areas as single area and termed it motor-somatosensory cortex (**M1C/S1C**) due to the lack of clear anatomical and histological boundaries between immature M1C and S1C. After the appearance of the central sulcus M1C was sampled in front of the central sulcus from the middle and upper part of the lower third of the dorsolateral surface of the hemisphere.

For ***periods 8–15***, sampled areas were as follows (Supplementary Fig. 1):

- OFC was sampled from the anterolateral two thirds of the orbital gyri. OFC corresponds approximately to BA 11.
- DFC was sampled from approximate border between the anterior and middle third of the medial frontal gyrus. DFC corresponds approximately to BA 9 and 46.
- VFC was sampled from the posterior third of the inferior frontal gyrus, corresponding to the opercular and triangular part of the inferior frontal gyrus. VFC corresponds approximately to BA 44 and 45.
- MFC was sampled from perigenual and subgenual parts of the anterior cingulate gyrus and the anteromedial part of the superior frontal gyrus. MFC corresponds approximately to BA 24, 32 and 33.
- M1C was sampled from the ventrolateral part of the precentral gyrus, corresponding most closely to the orofacial region of M1C. M1C corresponds to BA4.

2.2.1.2. Parietal cortex (PC)

For ***period 1***, the sampled areas corresponded to the dorsal middle third of the cerebral wall.

For ***period 2***, the sampled areas included the paracentral region corresponding approximately to the prospective motor and parietal somatosensory (M1C/S1C) cerebral wall, and the posterior half of the dorsal middle third of the cerebral wall corresponding approximately to the prospective inferior parietal cortex (IPC).

For ***periods 3–7***, prior to the appearance of gyri and sulci, multiple areas of the PC were sampled as follows (Supplementary Fig. 2):

- **Primary somatosensory cortex (S1C)**, prior to the appearance of the central sulcus, was sampled immediately caudal to the M1C (see M1C description above). After the appearance of the central sulcus, S1C was sampled behind the central sulcus from the middle and upper part of the lower third of the dorsolateral surface of the cerebral hemisphere adjacent to the M1C area.
- **Posterior inferior parietal cortex (IPC)** was sampled from the lower posterior part of the dorsolateral surface of the middle third of the cerebral hemisphere adjacent to the end of the lateral sulcus.

For ***periods 8–15***, sampled areas were as follows (Supplementary Fig. 1):

- S1C was sampled from the ventrolateral part of the postcentral gyrus adjacent to the M1C area. S1C corresponds to BA 1–3.
- IPC was sampled from the posterior half of the supramarginal gyrus. IPC corresponds approximately to BA 40.

2.2.1.3. Temporal cortex (TC)

For ***period 1***, the sampled area corresponded to the anterior two thirds of the lateral part of the posterior third of the cerebral wall.

For ***period 2***, the sampled areas included the posterior two thirds of TC corresponding approximately to the prospective auditory and superior temporal cortex (A1C/STC) cerebral wall, and the anterior third corresponding approximately to the prospective inferior temporal cortex (ITC).

For ***periods 3–7***, prior to the appearance of gyri and sulci, multiple areas of the TC were sampled as follows (Supplementary Fig. 2):

- **Primary auditory cortex (A1C)** was sampled from the upper part of the temporal bank of the lateral sulcus.
- **Posterior superior temporal cortex (STC)** was sampled from the upper part of the superior third of the temporal lobe adjacent to the lateral sulcus and A1C area.
- **Inferior temporal cortex (ITC)** was sampled from the lower part of the inferior third of the temporal lobe adjacent to the temporal lobe pole.

For ***periods 8–15***, sampled areas were as follows (Supplementary Fig. 1):

- A1C was sampled from the planum temporale and the transverse temporal gyri. A1C corresponds to BA 41.
- STC was sampled from the posterior third of the superior temporal gyrus. STC corresponds approximately to BA 22.
- ITC was sampled from the anterior third of the inferior temporal gyrus. ITC corresponds approximately to BA 20.

2.2.1.4. Occipital cortex (OC)

For ***periods 1 and 2***, sampled tissue corresponded to the posterior (occipital) part of the cerebral wall.

For ***periods 3–7***, (Supplementary Fig. 2) prior to the appearance of gyri and sulci, sampled tissue corresponded to prospective **primary visual cortex** (**V1C**). Prior to the appearance of the calcarine fissure, V1C was sampled from the posterior third of the medial wall of the prospective occipital lobe. After appearance of the calcarine fissure, V1C was sampled as described below.

For ***periods 8–15***, (Supplementary Fig. 1) V1C was sampled from the area surrounding the calcarine fissure. Only samples in which the stria of Gennari could be recognized were included. V1C corresponds to BA 17. Small pieces of the neighbouring BA18 could have been occasionally present in the sample, but the majority of the sample corresponded to BA17.

2.2.2. Hippocampus (**HIP**)

For ***periods 1 and 2***, HIP was sampled from the hippocampal anlage, located on the ventromedial side of the cerebral hemisphere.

For ***periods 3–15***, (Supplementary Figs 1 and 2) HIP was sampled from the middle third of the retrocommissural hippocampal formation, located on the medial side of the temporal lobe. Sampled areas always contained dentate gyrus and the cornu ammonis. Samples dissected from the frozen tissue may contain small quantities of the neighbouring choroid plexus.

2.2.3. Amygdala (AMY)

We aimed at dissecting the whole AMY from ***period 3–15*** specimens (Supplementary Figs 1 and 2). Very small quantities of surrounding white matter and potentially other surrounding structures in the basal telencephalon were included in samples.

2.2.4. Ventral forebrain (VF)

Depending on the time period we sampled different parts of the VF. For ***period 1***, the sampled region corresponded to the ventral forebrain (VF), which included primordium of the ganglionic eminence.

2.2.4.1. Medial ganglionic eminence (MGE)

2.2.4.2. Lateral ganglionic eminence (LGE)

2.2.4.3. Caudal ganglionic eminence (CGE)

We sampled the MGE, LGE and CGE separately from ***period 2*** specimens. Small quantities of surrounding tissue may be included in the samples.

2.2.4.4. Striatum (STR)

Striatum (**STR**) was sampled from ***periods 3–15*** specimens (Supplementary Figs 1 and 2). We dissected the anterior part of striatum containing the head of the caudate nucleus and the putamen, separated by the internal capsule and ventrally connected to the nucleus accumbens. Small quantities of surrounding white matter are included in the samples.

2.2.5. Diencephalon (DIE)

Depending on the time period, we sampled different parts of the DIE. For ***period 1***, the sampled region corresponded to the entire DIE.

2.2.5.1. Dorsal thalamus (DTH)

For ***period 2***, the sampled region corresponds to the dorsal part of the thalamic anlage (**DTH**).

2.2.5.2. Mediodorsal nucleus of the thalamus (MD)

For ***periods 3–15*** (Supplementary Figs 1 and 2), the whole mediodorsal nucleus of the thalamus (**MD**) was sampled from the dorsal and medial thalamus. Small quantities of surrounding thalamic nuclei could be present in the samples.

2.2.6. Upper (rostral) rhombic lip (URL)

Sampled region corresponds to the **URL** and adjacent tissue located above the upper rhomboid fossa for ***periods 1 and 2***.

2.2.6.1. Cerebellar cortex (CBC)

CBC was sampled from the lateral part of the posterior lobe for ***periods 3–15*** (Supplementary Figs 1 and 2). The sampled area contained all three layers of cerebellar cortex and underlying white matter but not the deep cerebellar nuclei. CBC approximately corresponds to the lateral pontocerebellum.

2.3. Workflow of quality control measures

To summarize the quality control (QC) measures taken in this study, we have made a diagram (Supplementary Fig. 5) showing QC steps during the generation of the dataset. Five QC steps were performed on 94 brains that were considered for this study. The first QC step was assessment of donors and tissue (see exclusion criteria in Supplementary section 3.3). Twenty-three brains did not pass the first QC step and were excluded from

further analysis. Of the remaining 71 brains that were included, 58 of them were subsequently processed (as of December 2010) and remaining brains are currently being processed. The second QC step was assessment of dissection process based on dissection score (DS; for explanation see Supplementary section 3.4.3). There were 74 dissected samples that did not pass this QC step, leaving 1,414 samples from 58 brains that were processed and from which RNA extracted for further analysis. The third QC step was control of RNA quality based on RNA integrity number (RIN; see Supplementary section 5.1), of which 15 samples did not pass. The fourth QC step was assessment of preparation and hybridization of RNA samples to the Affymetrix microarray chip (see Supplementary section 5.3). Following the removal of 53 samples after this QC step, 1,346 samples were successfully processed on Affymetrix microarray chips and analyzed for gene expression. The fifth QC step was detection of outliers. There were five samples, all from a single brain, excluded on this basis of hierarchical sample clustering (see Supplementary section 5.4, paragraph 1). One additional sample was excluded by correlation analysis (see Supplementary section 5.4, paragraph 2) As of December 2010, 1,340 microarray samples passed QC steps and were included in this study.

3. Tissue Procurement and Sampling

3.1. Tissue procurement

This study was conducted using postmortem human brain specimens from tissue collections at the Department of Neurobiology at Yale University School of Medicine and the Clinical Brain Disorders Branch of the National Institute of Mental Health. Additional specimens were procured from the Human Fetal Tissue Repository at the Albert Einstein College of Medicine (AECOM), the Brain and Tissue Bank for Developmental Disorders at the University of Maryland, the Birth Defects Research Laboratory at the University of Washington, Advanced Bioscience Resources Inc. and the MRC-Wellcome Trust Human Developmental Biology Resource at the Institute of Human Genetics, University of Newcastle, UK. Tissue was collected after obtaining parental or next of kin consent and with approval by the institutional review boards at the Yale University School of Medicine, the National Institutes of Health, and at each institution from which tissue specimens were obtained. Tissue was handled in accordance with ethical guidelines and regulations for the research use of human brain tissue set forth by the NIH (http://bioethics.od.nih.gov/humantissue.html) and the WMA Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html). Appropriate informed consent was obtained and all available non-identifying information was recorded for each specimen. Specimens range in age from 5.7 post-conceptual weeks (PCW) to 82 years. Of 57 postmortem brain specimens included in this study, 18 were obtained with either left

or right hemisphere, and 39 were obtained with both hemispheres. Embryonic and fetal age was extrapolated based on the date of the mother's last menstruation, characteristics of the fetus noted upon ultrasonography scanning, foot length of the fetus, and visual inspection. The postmortem interval (PMI) was defined as hours between time of death and time when tissue samples were frozen.

3.2. Neuropathological evaluation

All clinical histories, tissue specimens, and histological sections were evaluated to assess for hypoxia, cerebrovascular incidents, tumours, microbial infections, neurodegeneration, demyelination, and metabolic disease. In addition, cadavers from period 4 onward underwent a complete autopsy and were refrigerated beforehand to minimize degradation. To prepare tissue sections for microscopic neuropathological histological examination, small samples (usually the dorsal parietal cortex, striatum with ependymal layer and subependymal zone, hippocampus, and cerebellar cortex) of fresh or frozen tissue were dissected and fixed in 4% paraformaldehyde and processed for histology and immunohistochemistry as described below.

3.3. Selection criteria for brain specimens

To better ensure consistency between samples and decrease potential variation due to ante- and postmortem conditions, specific selection criteria were arbitrarily established. Most postnatal and adult brains also underwent comprehensive toxicological screening. The aim was to collect tissue specimens from clinically unremarkable donors without history or signs of neurological or neuropsychiatric illness or drug abuse. The following selection criteria were strictly adhered to when deciding whether to exclude or include each brain specimen.

- Brains with chromosomal or large-scale genomic abnormalities, detected by karyogram and/or Illumina Human Omni-2.5, were excluded.
- Prenatal and neonatal specimens were excluded if drug or alcohol abuse by the mother during pregnancy was reported or if potassium chloride, salt water, or urea were injected into the amniotic sac during surgical procedure.
- Only brains free of obvious malformations or lesions were collected. Disqualifying characteristics included any obvious abnormality of the neural tube, forebrain, brainstem, cranial nerves, cerebellum, or spinal cord (i.e., prominent intraparenchymal haemorrhage and ischemia, infection, periventricular leukomalacia, abnormal meninges, dysplasia, hypoplasia, alterations in the pial or ventricular surface, extensive white matter lesions).
- Samples were excluded if microscopic analysis revealed extensive neuronal loss, neuronal swelling, glioneuronal heterotopias, or dysmorphic neurons and neurites.

- Samples that tested positive for Hepatitis B, Hepatitis C, or HIV were excluded.
- Early postnatal and adult (periods 8–15) specimens were excluded if excessive drug or alcohol abuse was reported, if the individual had any known neurological or psychiatric disorders, or if any prolonged agonal conditions (coma, prolonged pyrexia, hypoxia, seizures, prolonged dehydration, hypoglycaemia, multiple organ failure) were reported. Other excluding factors included ingestion of neurotoxic substances at the time of death, suicide, severe head injury, significant haemorrhages, widespread vascular abnormalities, ischemia, tumours, stroke, congenital neural abnormalities, and signs of neurodegeneration (spongiosis, amyloid plaques, neurofibrillary tangles, Lewy bodies, and amyloid angiopathy).

3.4. Tissue processing and dissections

Depending on the condition and period of the procured specimens, four different dissection methods were used. Photos and/or video were used to document dissections using a digital camera. Regions of interest were matched between different specimens, ages, and hemispheres of each brain. Supplementary Table 3 provides a complete list of all tissue samples for which we have included Exon array data in this study. Specific dissection protocol depended upon the period of the specimen and the method by which it was preserved. For all brain specimens procured at Yale University School of Medicine and the Human Fetal Tissue Repository at AECOM, brain regions and NCX areas of interest (Supplementary Table 1) were collected from fresh tissue. For all other specimens, regions/areas were collected from frozen tissue slabs or whole specimens stored at -80 ºC. To ensure consistency between specimens, all dissections were performed by the same person. Small samples of fresh or frozen URL (period 1 and 2) or CBC (period 3-15) were used to measure tissue pH.

3.4.1. Tissue dissection methods

Different dissection procedures were used for each specimen, depending upon the period of the brain (see below). Our pilot experiments indicated that the quality of RNA and DNA was largely unaffected by variation between the dissection methods used.

3.4.1.1. Regional sampling from fresh brain specimens

Brains were chilled on ice for 15–30 minutes prior to sectioning. Brains were placed ventral side up onto a chilled aluminium plate (1 cm thick) on ice. The brainstem and cerebellum were removed from the cerebrum by making a transverse cut at the junction between the diencephalon and midbrain. Next, the cerebrum was divided into left and right hemispheres by cutting along the midline using a Tissue-Tek Accu-Edge trimming blade,

260 mm. The cerebellum was separated from the brainstem by cutting directly posterior to the brainstem, along the cerebellar peduncles. The regions of interest were dissected using a scalpel blade and immediately frozen in liquid nitrogen. Dissected samples were either immediately processed for RNA extraction or stored at -80 °C for later RNA extraction. The remaining brain tissue was cut to obtain 1 cm (period 5–15 specimens) or 0.5 cm (periods 3 and 4 specimens) thick serial, coronal sections. The tissue slabs were snap frozen in isopentane/dry ice at -30 to -40 ºC and stored at -80 ºC.

3.4.1.2. Regional sampling from frozen brain specimens

All previously frozen periods 3–15 specimens and tissue slabs were microscopically inspected and the desired region was demarcated, then dissected using a dental drill (AnyXing, 300D) and a Lindemann Bone Cutter H162A.11.016 or diamond disk saw (Dental Burs USA; r=11 mm) on a 1 cm thick aluminium plate over dry ice. Dissected tissue samples were stored at -80 ºC prior to further processing.

3.4.1.3. Regional sampling from specimens processed in RNA*later*

Frozen period 2 specimens were sectioned coronally at approximately 2 mm, beginning at the frontal pole, using a dental diamond disk saw. For gradual thawing, tissue slabs were transferred from -80 °C storage to overnight storage in RNA*later* ICE (Ambion) at -20 °C. Tissue slabs were visually inspected for gross anatomical neuropathological abnormalities. Next, regions of interest were sampled under a dissection microscope at 4 °C and stored in Buffer RLT Plus from the RNeasy Plus Mini Kit (Qiagen) at 4 °C. RNA was immediately extracted.

3.4.2. Histological verification of tissue sampling

To verify that the region or NCX area of interest is properly and consistently sampled, we also collected small tissue blocks, from both frozen and fresh brain specimens, adjacent to the tissue sample dissected for the RNA extraction. We have done this for the majority of M1C, S1C, IPC, A1C and V1C samples, which in our experience were hard to match across different specimens but can be histologically verified using Nissl method in postnatal specimens due to cytoarchitectonic differences. This method was also occasionally used for other regions or NCX areas. These tissue blocks were then fixed in 4% paraformaldehyde for 48 h, sectioned at 50 µm thickness using a vibratome, and Nissl stained to verify the identities of dissected adjacent tissue (examples of Nissl stained brain regions and NCX areas from a fetal and adult brain are provided in Supplementary Fig. 3).

3.4.3. Dissection scoring

We developed a scoring system to evaluate the precision of how well the sampled region/area was represented at the same position of corresponding samples of the same period.

Score	Sample description
1 or **2**	The region/area of interest was absent (1) or largely absent (2) and thus not collected.
3	The region/area of interest was not complete but was of suitable quality to collect.
4	The region/area of interest was largely intact but was not histologically verified or could not be collected at precisely the same position from which the corresponding contralateral sample was collected.
5	The region/area of interest was fully intact, verified by gross inspection or Nissl staining (NCX areas), and collected at precisely the same position as corresponding samples of the same period.

3.4.4. Tissue pulverization. To ensure proper representation of the region of interest, frozen tissue samples were pulverized in liquid nitrogen using a ceramic mortar and pestle (Fisher Scientific, cat# 12-961C and 12-961-5C). Pulverized samples were transferred to wide-mouth cryogenic vials (Nalgene, cat# 03-337-7B) and stored at -80 ºC until used for RNA extraction.

4. DNA Isolation and Genotyping Data Analyses

4.1. DNA extraction and genotyping

For genotyping analysis, up to 25 mg of brain tissue, usually collected from the CBC, was homogenized using a bead mill homogenizer (Bullet Blender, Next Advance) and lysed in Buffer ATL (Qiagen) at 56 °C for 3–4 h. Genomic DNA was isolated using a non-phenolic procedure (DNeasy Blood & Tissue Kit, Qiagen) followed by proteinase K and RNase A treatment (Qiagen). Optical density values of extracted DNA were measured using a NanoDrop (Thermo Scientific) and PicoGreen dsDNA assay kit (Invitrogen). DNA integrity was confirmed by agarose gel electrophoresis. Illumina Omni-2.5 million SNP arrays were used for genotyping. DNA samples were processed at the Yale Center for Genome Analysis according to the Infinium HD Assay Super, Automated Protocol for Human Omni 2.5-Quad Bead Chip (Illumina).

4.2. Copy number variation (CNV) analysis and genomic quality control

All SNP chips were scanned using the Illumina iScan system. The intensity files were analysed using Illumina GenomeStudio v2010.2 software. Sex of each sample was determined from the SNP genotyping results using GenomeStudio to confirm the metadata of each sample. Two algorithms were used to detect CNV from SNP intensity data. The cnvPartition algorithm in GenomeStudio was used to detect CNV. The measurement of B allele frequency and log-transformed R ratio of all SNPs were exported to the program PennCNV to confirm CNVs. The two sets of results were compared and only CNVs detected by both algorithms were included in the final results (Supplementary Fig. 4). Only specimens with no signs of chromosomal or large-scale genomic abnormalities were considered for the study.

4.3. Corroborating and refining ethnic background from genotypes

Ethnicity of donors was reported by next of kin. However, in the majority of prenatal and some postnatal cases the ethnicity of the father was not available. For this reason, the ethnic background was corroborated and refined by comparing SNP analysis of each individual to SNP data available in HapMap III. The whole allele frequency dataset of the following ethnic populations was downloaded from HapMap III (ftp://ftp.ncbi.nlm.nih.gov/hapmap/frequencies/2010-08_phaseII+III/): (1) Utah residents of Eastern and Western European descent (CEU); (2) Toscani in Italy (TSI); (3) Yoruba, in Ibadan, Nigeria (YRI); (4) African ancestry in Southwest USA (ASW); (5) Luhya in Webuye, Kenya (LWK); (6) Maasai in Kinyawa, Kenya (MKK); (7) Chinese in Metropolitan Denver, Colorado (CHD); (8) Japanese in Tokyo, Japan (JPT); (9) Mexican ancestry in Los Angeles, California (MEX) and (10) Gujarati Indians in Houston, Texas (GIH).

A group of SNPs frequently expressed in some populations but rarely in others was first selected. The maximum allele frequency and minimum allele frequency in the 10 populations were calculated for each SNP. Population specific SNPs were defined as those for which the difference between the maximum and minimum allele frequency was greater than 0.6. These population-specific SNPs were cross-referenced against the SNPs identified by genotyping of our samples. Only the common set of SNPs was used in the following calculation.

The common set of SNPs is designated as S_i, where i is a variable value that represents a specific SNP. The genotype G_i of SNP S_i could be AA, AB or BB, where A and B are two alleles for the SNP. The populations are designated as P_j, where j is a value of 1 through 10 that represents a specific ethnicity, as defined above. The A-allele frequency of SNP S_i

within population P_j is f_{ij}. For any individual in a specific population P_j, the probability of having the genotype G_i at SNP S_i is:

$$p_{ij} = \begin{cases} f_{ij} * f_{ij} & if\ G_i = AA \\ 2 * f_{ij} * (1 - f_{ij}) & if\ G_i = AB \\ (1 - f_{ij}) * (1 - f_{ij}) & if\ G_i = BB \end{cases}$$

The log-transformed likelihood function (L_j) that describes the probability of any individual from population P_j having exactly the same genotypes as the analyzed samples at a specific set of SNPs would be:

$$log_{10}(L_j) = log_{10}\left|\prod_{i=1}^{n} p_{ij}\right|$$

$$= \sum_{i=1}^{n} log_{10}(p_{ij}).$$

Therefore, the log-transformed likelihood ratio for population P_j compared to the reference population P_0 is:

$$log_{10}(L_j/L_0) = \sum_{i=1}^{n}\left[log_{10}(p_{ij}) - log_{10}(p_{i0})\right]$$

The likelihood ratio for each population was used to assign ethnicity. “European Ancestry” was assigned to samples for which the two highest-score populations were CEU and TSI. “African Ancestry” was assigned if the four highest-score populations were YRI, ASW, LWK and MKK. “Asian Ancestry” was assigned if the three highest-score populations were JPT, CHD, and GIH. “Hispanic Ancestry” was assigned if the highest-score population was MEX. “Mixed” was assigned as the default ancestry if none of the above conditions could be met.

5. RNA Isolation, Exon Array Processing, and Quality Assessment

5.1. RNA extraction

A bead mill homogenizer (Bullet Blender, Next Advance) was used to lyse the pulverized tissue. Each pulverized tissue sample was transferred to a chilled safe-lock microcentrifuge tube (Eppendorf). A mass of chilled stainless steel beads (Next Advance, cat# SSB14B) equal to the mass of the tissue was added to the tube. Two volumes of Buffer RLT Plus (Qiagen) were added to the tissue and beads. Samples were mixed in the

Bullet Blender for 1 min at a speed of six. Samples were visually inspected to confirm desired homogenization and then incubated at 37 °C for 5 min. Buffer RLT Plus was added up to 0.6 ml, and samples were mixed in the Bullet Blender for 1 min. Total RNA was extracted using a non-phenolic procedure (RNeasy Plus Mini Kit, Qiagen), followed by DNase treatment (TURBO DNase, Ambion) as per manufacturers' instructions. Optical density values of extracted RNA were measured using a NanoDrop (Thermo Scientific) to confirm an A_{260}:A_{280} ratio above 1.9. RIN was determined for each sample using Bioanalyzer RNA 6000 Nano Kit or Bioanalyzer RNA 6000 Pico Kit (Agilent), depending upon the total amount of RNA.

5.2. Exon array hybridization

Exon array hybridizations were performed at the Yale Center for Genome Analysis and at Gene Logic Inc. (Gaithersburg, MD). Reverse transcription (RT) was performed to generate cDNA from total RNA using RT primers designed using an oligonucleotide matching algorithm. For the selective cDNA synthesis, the Ambion WT Expression kit (Ambion) was used in combination with the GenechipWT Terminal Labeling and Controls Kit (Affymetrix) for target preparation, according to manufacturer recommendations. PolyA controls were added to the input RNA to measure efficiency of target amplification. Fragmented and labelled second cycle cDNA (5.5 µg) was added to a hybridization cocktail prior to loading of 200 µl onto individual Affymetrix Human Exon 1.0 ST arrays, which features comprehensive coverage of the human genome, with 1.4 million probe sets that assay exon expression across the entire transcript. Microarrays were hybridized at 45 °C for 16–24 hours, washed and stained using an Affymetrix FS450 fluidics station, according to manufacturer recommendations. Microarrays were scanned on a GeneChip Scanner 3000 and visually inspected for hybridization artifacts. Exon chip analysis was performed using Affymetrix Power Tools 1.12.0. Probe level data was summarized into probe set level data using the Robust Multichip Average (RMA) background correction algorithm in combination with an R-script. The raw image files (.DAT files) were analysed using Affymetrix GeneChip Operating Software to generate .CEL files.

5.3. Exon array quality assessments

Several quality control measures were implemented throughout microarray sample preparation and transcriptome data generation steps (Supplementary Fig. 5) to reduce errors due to spatial artifacts on the chips, technical differences between chips in probe saturation, differences in the intensity of the probes along the 5'to 3' gradient of genes, or due to other unaccounted batch effects. First, an idealized reference chip was constructed for all arrays hybridized at each facility by computing the 15% trimmed mean of the $\log_2$-

transformed probe intensities for each of the 5.5 million perfect match probes across all arrays. Spatial artifacts were defined as severe non-random spatial patterns in the ratios of intensities of probes on one chip relative to corresponding probes on the ideal reference chip; such artifacts are believed to arise from non-uniform hybridization conditions across the surface of an array. By design, microarray probes are randomly distributed across an array to avoid spatial biases, i.e., probes that represent adjacent regions on the genome are not located physically adjacent on an array. Consequently, in the absence of spatial artifacts due to error sources in pre-processing steps, one would expect a random pattern of probes with higher or lower intensities relative to the reference chip. Supplementary Figure 6 shows ratio intensity plots for arrays with severe spatial artifacts.

In addition, exon array hybridization uniformity is estimated by gene expression uniformity from 5'-end to 3'-end (Supplementary Fig. 6c). Samples with altered hybridization uniformity and microarrays displaying regional biases or other spatial artifacts were excluded/or re-tested. Low-quality RNA samples identified by these quality control measures were excluded from further analysis (N=53).

To evaluate technical reproducibility and the possibility of batch effects, four identical samples were submitted to and processed at different testing centers. To evaluate biological reproducibility, we calculated the correlations between the same regions collected from different individuals of the same period. The correlations were high for both technical (Spearman correlation, r^2 = 0.967; N=8 total) and biological replicates (r^2 = 0.926; 1,340 total), reinforcing the validity of our approach and data.

The impact of confounding factors on the quantity, quality, and transcriptional profile of RNA was controlled for by computing correlations between gene expression and the following variables: PMI, pH, RIN, and DS. Of these factors, PMI and RIN showed a weak anticorrelation and correlation, respectively, with the number of expressed genes (Spearman correlation, r = -0.382 and 0.278, respectively; Supplementary Fig. 12), indicating that the collected tissue samples are suitable for profiling the transcriptome.

5.4. Detection of outliers

We used several approaches to identify outliers based on processed exon array samples. First, we performed average-linkage hierarchical clustering of all samples corresponding to different regions using 1-correlation as a distance and related them to the following variables: age, sex, ethnicity, PMI, DS, RIN, and the array processing site. These confounding factors for each sample were labeled by color. For the NCX, the expression and the RIN of areas from the same hemisphere of an individual were averaged. Hierarchical clustering revealed notable clustering of the samples based on age and regions (Supplementary Fig. 8). Out of 1,346 samples, 5 samples from the same brain

were found to be outliers based on hierarchical clustering and subsequently removed from further analyses.

Second, the averaged Spearman's correlation coefficient of a region/area was calculated for a period, and those samples with values outside 3 standard deviations (SDs) of the mean were considered as outliers. Individual samples or brain specimens identified as outliers by either hierarchical clustering (N=1 brain, 5 samples total) or correlation analysis (N=1 sample) were re-processed and/or excluded from the study. Finally, a total of 1,340 exon microarray samples from 57 brains passed the quality control steps and were included in this study (Supplementary Tables 3 and 4).

5.5. Exon array data pre-processing

Partek Genomics Suite version 6.5 (Partek Incorporated, St. Louis, MO, USA) was used to normalize raw exon array data and to summarize expression of the probe set and transcript cluster. Affymetrix CEL files that passed QC analyses were imported into Partek Genomics Suite using the default Partek settings: RMA background correction[58], quantile normalization, mean probe set summarization, and $\log_2$-transformation. All analyses reported used only the core and unique probe sets, which hybridize to unique genomic positions and are based on well-curated RefSeq (www.ncbi.nlm.nih.gov) and Ensembl transcripts (www.ensembl.org) as well as GenBank transcripts (http://www.ncbi.nlm.nih.gov/genbank/) annotated as complete protein-coding sequences. A total of 17,565 protein-coding genes were surveyed.

Datasets were annotated according to the UCSC human genome hg19 reference sequence (http://genome.ucsc.edu/cgi-bin/hgGateway). 105,271 core probes (within 62,448 probe sets out of 230,918 core probe sets) contained SNPs defined in the probe group file HuEx-1_0-st-v2.r2-SNPs-Excluded.pgf provided by Affymetrix, which is based on the dbSNP database (version 129, April 2008) and SNPinprobe_1.0 database. As previously observed[59,60], inclusion of SNP-containing probes can affect the hybridization of the probes and thus produce misleading signal intensity values. To test the effects of these SNP-containing probes, the gene and exon expression were evaluated twice, with one evaluation excluding SNP-containing probe sets by filtering out the affected probes that hybridized to regions with common SNPs (minor allele frequency (MAF) ≥ 0.05). The Pearson correlation analysis was performed to compare gene and exon expression levels of the data with and without SNP-containing probe sets. Almost all of the correlation coefficients (98.8%) for all 1,340 samples were within the range of 0.980-0.995 for genes and 0.975-0.995 for exons, suggesting that SNP-containing probe sets have a very minimal effect on measured expression levels (Supplementary Fig. 7). Nevertheless, we removed SNP-containing probe sets during the normalization step in the Partek program

to be control for SNP-related confounding effects. The median of all individual probe sets of one transcript cluster was used as the estimate of gene expression values. To filter out low expression signals (including noise or poorly hybridized probes), which may lead to false positives, detection above background (DABG) *P*-values of exon probe sets and transcript clusters were calculated using Affymetrix power tools (APT, http://www.affymetrix.com/partners_programs/programs/developer/tools/powertools.affx).

6. Exon Array Data Analysis

6.1. Gene-level analysis

The expression level for all core and unique probe sets within an exon were averaged to obtain an expression value for the exon, and the median of all exons within one gene (transcript cluster) was used as the estimate of gene expression. To reduce the noise in the dataset and the false positives in the following differentially expressed genes and alternative exon usage investigations, we excluded genes with a $\log_2$-transformed expression value <6 in all 1,340 samples. We defined an "expressed" gene as a gene with a $\log_2$-transformed expression value ≥6 in at least one sample and mean detection above background (DABG) P<0.01 in one or more of the sampled brain regions or NCX areas during at least one time period. To demonstrate the underlying distribution of the data using these criteria to define expression, the signal intensity distribution plots of expressed, non-expressed gene, and intronic controls (designed from the intron regions of a set of housekeeping genes) are shown in Supplementary Figure 9.

6.2. Multidimensional scaling (MDS)

MDS analysis was performed by generating expression vector for all expressed genes for each of the analyzed samples. We calculated the Euclidean distance between each pair of vectors to obtain a distance matrix with dimension 1,340 x 1,340. The R function *isoMDS* was used to calculate two-dimension MDS. The returned two vectors were used as coordinates x and y for each of the 1,340 samples to make a scatter plot in a 2-dimension plane. The 1,340 data points, each representing one tissue/array sample, were plotted and colored according to the phenotype of the samples, such as time period, brain region, sex, ethnicity, or donor's code (Supplementary Fig. 10).

6.3. Principal component analysis (PCA)

PCA from prcomp package and three-dimensional view from rgl package, both retrieved from R (http://www.r-project.org/), were used to identify and visualize the major contributors. They were both used to analyze the gene expression of 11 neocortical areas and 6 regions across 15 time periods from all 1,340 samples separately. The major

contributors will dominate the clustering of samples. Our results show that the samples were clustered by developmental periods and functional regions. Three different directional views are shown in order to visualize the clustering (Supplementary Fig. 11).

6.4. Identification of spatial and temporal DEX genes

Analysis of variance (ANOVA) was used to identify differentially expressed (DEX) genes across all regions of interest and during all periods (1-15). Genes that were differentially expressed in at least one brain region or NCX area were identified as spatial DEX based on ANOVA by using a region/area as an ANOVA factor. Genes differentially expressed in at least one time period were identified as temporal DEX based on ANOVA using a period as a factor. In order to exclude the possibility that variation in RIN scores and PMI within the acceptable range might influence procedures, we included PMI and RIN as technical covariates within our ANOVA model of differential expression. We found that RIN accounted for a mean variance of 2.16 in the spatial analysis for all genes, which was 8% of the mean total variance across genes, and a variance of 1.02 in the temporal analysis, which was 3% of the mean total variance of that analysis. We found that PMI accounted for a mean variance of 1.89 in the spatial analysis, which was 8% of the mean total variance, and a variance of 0.33 in the temporal analysis, which was 1% of the mean total variance of that analysis. Resulting *P*-values from ANOVA were corrected for multiple comparisons using the Benjamini and Hochberg false discovery rate (FDR) method[62]. A conservative statistical threshold (FDR <0.01 and minimum fold difference >2 between brain regions/areas or periods) was used to identify DEX genes. Genes that were not significantly expressed above background were excluded from ANOVA tests.

To ensure that our results are robust and are not unduly influenced by any particular sample, we used a five-fold jackknife procedure. Samples were split into 5 random groups and re-analyzed for DEX genes 5 times by leaving out one group in each cycle of analysis. The genes from each of the five re-sampling showed a high rate of overlap with the genes obtained from the whole dataset (Supplementary Fig.13).

To identify spatially DEX genes that were selectively highly enriched or restricted in only one of the regions, irrespective of their temporal regulation, we used a combination of t-test and fold difference analyses. We limited this spatial DEX analysis to periods 3-15, when regions/areas of interest are defined using equivalent criteria and can be consistently followed across time. We used a sliding-window algorithm to detect regionally enriched genes within a group of sequential periods. The window size was set as 5, scanning periods 3-7, 4-8, 5-9, 6-10, 7-11, 8-12, 9-13, 10-14, 11-15. For each window, a two-sample t-test was applied to determine if the expression level of a gene in one selected region was significantly different from the expression level in the samples from

other regions. To calculate the fold difference in averaged signal intensity, two regions were used: the region of interest and, from the remaining regions, the one that had the highest averaged expression value. The P-values from the t-tests were transformed to FDR using the Benjamini and Hochberg method. A gene with greater than 2-fold differences in a given window was considered to be highly enriched in or restricted to one region. If the window contained both prenatal and postnatal periods or both postnatal and adult periods, a final step was performed in order to determine if the DEX genes should be assigned to one or both of the developmental groups. We calculated the fold difference attributed to the two periods bordering each side of the developmental groups (6-7 and 8-9 for prenatal-postnatal group; 11-12 and 13-14 for postnatal adult group). If the fold difference was greater than 1.5, the DEX gene was assigned to that developmental group, resulting in either one or both groups being assigned the gene.

6.5. Analysis of DEU

To identify expressed genes that exhibit differential exon usage (DEU) across all regions of interest and during all periods, we used the splicing ANOVA method and the splicing index (SI) algorithm[61]. Normalized intensities (NI) were calculated as the expression of an individual exon relative to the expression of the gene. Splicing index was defined as the fold difference of the normalized intensity of exons between two groups (periods and/or regions). Both spatial DEU and temporal DEU were analysed. For spatial DEU, the ANOVA factor is set as the brain region/area. All periods were tested and the maximum of splicing indexes among regions was calculated. For temporal DEU, the ANOVA factor is set as the period. Sixteen regions/areas were tested, and the maximum of splicing indexes among periods was calculated. We included PMI and RIN as technical covariates within our ANOVA models testing alternative exon usage. Resulting *P*-values from ANOVA were corrected for multiple comparisons using the Benjamini and Hochberg FDR method[62]. To detect meaningful DEU, stringent criteria (FDR < 0.01 and maximum splicing index > 2) were used, and exons expressed at low levels or with low variance (standard deviation of exon expression level <0.5 in all samples, which indicates they are very likely to be saturated in all samples) were excluded.

To test the robustness of our DEU, we used a five-fold jackknife procedure. Samples were split into 5 random groups and re-analyzed for DEU genes 5 times by leaving out one group in each cycle of analysis. The genes from each of the five re-sampling showed a high rate of overlap with the genes obtained from the whole dataset (Supplementary Fig.13). These analyses indicate that our findings are robust and did not depend on any particular sample.

6.6. Identification of sex-bias in DEX and DEU

Genes differentially expressed (DEX) between males and females were identified by the combination of a sliding window algorithm, two-sample t-test, and fold difference analysis. The sliding-window algorithm was used to detect sex-specific gene expression differences within a group of sequential periods. The window size was set as 5, which ensured adequate representation of both sexes, and periods 2-6, 3-7, 4-8, 5-9, 6-10, 7-11, 8-12, 9-13, 10-14, 11-15 were scanned. For each window, a two-sample t-test was applied to determine if the expression level of a gene in male samples was significantly different from the expression level in female samples in the same region. The *P*-values from the t-tests were transformed to FDR using the Benjamini and Hochberg method[62]. For each gene, the fold difference (log_2-transformed) between male and female samples in each region was also calculated. A FDR cutoff of 0.01 and 2-fold difference were used as a cutoff to identify genes that are DEX between males and females in each window. If the window contained both prenatal and postnatal periods or both postnatal and adult periods, a final step was performed in order to determine if the differentially expressed genes should be assigned to one or both of the developmental groups. We calculated the fold difference attributed to the two periods bordering each side of the developmental groups (6-7 and 8-9 for prenatal-postnatal group; 11-12 and 13-14 for postnatal-adult group). If the fold difference was greater than 1.5, the differentially expressed gene was assigned to that developmental group, resulting in either one or both groups being assigned the gene.

Sex-bias in differential exon usage (DEU) was investigated by splicing index and two-sample t-test. Each exon cluster is normalized by dividing its expression by the gene-level expression to yield a normalized intensity (NI). The splicing index is then calculated simply as a "fold change" between the normalized levels of each exon cluster from male and female samples. As in the sex-specific DEX analysis, the sliding window algorithm was also used. A FDR cutoff of 0.01 and 2-fold difference were used for sex-specific DEU. Exon clusters that were expressed at low levels (log_2 intensity <6) or with low variance (standard deviation of exon expression level <0.5 in all 1,340 samples, which indicates they are very likely to be saturated in all samples) were excluded. Y chromosome genes were also excluded from the analysis.

Analyses of sex-bias in DEX and DEU were performed without considering PMI and RIN as covariates because we did not find any significant differences of PMI and RIN values between female and male tissue samples. PMI, 12.01 ± 9.01 (N=31, mean ± SD) for males and 11.4 ± 8.09 (N=26, mean ± SD) for females; RIN, 8.86 ± 0.87 for males and 8.79 ± 0.99 for females.

6.7. Weighted gene co-expression network analysis

6.7.1. Dataset filtering

Only samples from periods 3–15 were included in the weighted gene co-expression network analysis (WGCNA). During these periods, brain regions and NCX areas are well-defined using equivalent criteria and can be consistently followed across time. NCX areas were not present in periods 1 and 2.

6.7.2. Network construction and module detection

To reduce noise, only genes with $\log_2$-intensity values greater than 6 in at least one sample and a coefficient of variance greater than 0.08 were used. A total of 9,093 genes fit these criteria. Signed weighted gene co-expression network analysis was performed using WGCNA R package[65]. General information about network analysis methodology and WGCNA software is available at www.genetics.ucla.edu/labs/horvath/CoexpressionNetwork. Pairwise Pearson correlation coefficients were calculated for all genes selected. The resulting Pearson correlation matrix was transformed into a matrix of connection strengths (an adjacency matrix) using a power function (connection strength = $(\frac{1+\text{correlation}}{2})^{\beta}$), which was then converted to a topological overlap matrix. WGCNA seeks to identify modules of densely interconnected genes by hierarchical clustering based on topological overlap [62]. The first principal component of each module was calculated and smoothed by smoothing spline against $\log_{10}$ (age in days). The top 50 genes expressing the highest within module connectivity were selected and exported to VisANT for network visualization[66].

6.7.3. Module filtering

The grey module consisted of genes not assigned to any of the module was removed. To check the effects of RIN, PMI, dissection, and pH on module identification, we constructed a linear model $PC1_i = age_i + region_i + sex_i + RIN_i + PMI_i + pH_i + DS_i$ (where $PC1_i$, age_i, $region_i$, sex_i, RIN_i, PMI_i, pH_i, DS_i, are the first principle component, $\log_{10}$ [days after conception], sex, RIN, PMI, pH and DS of sample i, respectively). ANOVA was then performed, and the coefficient of determination (r^2) and *P* value of each factor was calculated. The modules with r^2 of any factor greater than 0.1 and r^2 of any confounder (RIN, PMI, pH and DS) larger than the largest r^2 of the main factors (age, region and sex) were considered as modules corresponding to confounders. Five modules were detected as confounder-related modules and removed from further analysis.

6.7.4. Comparison of modules with transcriptional profiling of neurobiological categories

To gain insight into module functions, we compared each module with our manually created functional gene lists for specific neurobiological categories. The correlation between module eigengene (the first principle of module expression) and the PC1 summarization of each functional gene list was calculated and listed in Supplementary Table 10.

6.8. Gene ontology (GO) enrichment analysis

Functional enrichment was assessed using DAVID Bioinformatics Resources 6.7[63,64] (http://david.abcc.ncifcrf.gov/).

6.9. Creation of gene lists for and transcriptional profiling of neurobiological categories

We manually curated lists of genes functionally related to specific neurodevelopmental processes, neural cell types, and neurotransmitter systems, by selecting individual genes or a small group of genes from the existing gene ontology (GO) database (http://www.geneontology.org/) whose functions, and thus expressions, have been most closely associated with specific neurobiological category based on published data from model organisms and human medical genetics. We expected that the selected genes or gene group may accurately reflect the trajectories of neurobiological processes in neurodevelopment. To summarize the principle gene expression profile of each category, PCA was performed. The resulting first principal component (PC1) was plotted against logarithmic age in days, and a smooth curve was fitted by smoothing spline to display the developmental trajectories. For categories with only one gene, the expression level was used as PC1 (Supplementary Table 12 and data available on the website).

We also curated gene lists for neurotransmitter systems by compiling genes encoding critical enzymes for transmitter synthesis and receptor subunits. To compare the spatiotemporal patterns of genes in the same group, we plotted genes in each group individually and fitted the individual gene expression pattern by smoothing spline. Each group has 6 plots, one for each region (Supplementary Table 12 and data available on the website).

To test the accuracy of this strategy, three gene expression trajectories (*DCX*, dendrite development genes, synapse development genes) were compared with independently generated, non-transcriptome datasets for DCX-positive cell density, synaptic density, and the number of basal dendrites in the corresponding brain regions/areas, respectively

(Figure 5). The independent data were scaled by $\frac{x-\mu}{\sigma}$. μ and σ are the mean and the standard deviation of values, respectively, corresponding to the time periods for which both our gene expression and the independently generated data were available. Our scaled gene expression data were used to generate a spline curve, and the predicted values on this curve were calculated corresponding to available time points of previously generated independent data on DCX, dendrite development, and synaptogenesis, which used samples at different time points but within our defined periods. We then calculated the correlation coefficient between our predicted data and the actual values.

To compare the developmental trajectories for cell proliferation, neuron migration, dendrite development, myelination and synapse development gene markers, PC1 of each group was subjected to smoothing spline curve fitting. The fitted value $\hat{y}_i$ was then scaled to 0 to 1. The range of fitted vector $\hat{Y}$ is set to 1, and the scaled values were calculated as $z_i = (\hat{y}_i - \min(\hat{Y}))/\text{range}(\hat{Y})$. The scaled values were then plotted to represent the trajectory of each group with a measure of the percentage of changes.

To select genes potentially associated with the neurodevelopmental processes and neural cell types, the top 100 genes positively correlated with PC1 of each functional gene list were presented in Supplementary tables 13 and 14.

6.10. Expression trajectories and gene correlations for ASD- and schizophrenia-associated genes

A list of genes commonly associated with autism spectrum disorders (ASD) was obtained from recently published reviews[10,67] and relevant databases (www.g2conline.org). A list of genes commonly associated with schizophrenia was obtained from meta-analysis list[68] and relevant databases (www.szgene.org/topresults.asp; www.g2conline.org). In the szgene database list, genes with an assigned overall grade A, which indicates the highest/strictest association, were included in the analysis. Lists of genes for the two disorders and other correlated transcripts are provided in Supplementary Tables 15 and 16.

To generate representations of the dynamics in spatiotemporal expression of representative genes (*CNTNAP2, MET, NLGN4X, NRGN*), a heat map matrix was created for each of these genes according to the following method. For each combination of developmental period and brain region/area, the log2-transformed expression level values of related samples were averaged to obtain a single value of expression level at the specific period and specific region/area. All of these average values were collected into a data matrix, where each row represented one brain region/area and each column

represents one period. After the matrix was created, a heat map plot was created using R function 'heatmap.2' in package 'gplots' (Supplementary Figure 25).

The top 50 genes, ranked by their correlation with individual genes previously associated with ASD and schizophrenia, are presented in Supplementary Table 15 and 16. Genes with the highest correlation to the disease-related genes were grouped, and subjected to functional enrichment assessment using DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/).

7. Transcriptome Validation Methods

7.1. Quantitative real time RT-PCR and semi-quantitative RT-PCR

An aliquot of the total RNA that was previously extracted from each brain region was used for secondary validation through real-time PCR analysis. One µg of total RNA was used for cDNA synthesis using SuperScript III First-strand synthesis Supermix (Invitrogen) and subsequently diluted with nuclease-free water to 1 ng/µl cDNA. Gene-specific high-melting temperature primers for genes of interest were designed using NCBI/Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and expressed sequence information obtained from GenBank (NCBI). PCR reactions were conducted on an ABI 7900 Sequence Detection System (Applied Biosystems) using a hot start SYBR-green based method (Fast SYBR Green Master Mix, ABI) followed by melt curve analysis to verify specificity of the product. The Ct value (cycle number at threshold) was used to calculate the relative amount of mRNA molecules. The Ct value of each target gene was normalized by subtraction of the Ct value from housekeeping genes to obtain the ΔCt value. The relative gene expression level was shown as $2^{-\Delta Ct}$ or -ΔCt. All human genes of interest were normalized to the housekeeping gene *GAPDH.* Semi-quantitative RT-PCR was performed using the same template and hot start Taq DNA polymerase (Qiagen) under the following conditions: activation at 95 ºC for 10 min, followed by 30–40 cycles at 94 °C for 30 sec, 56 °C for 30 sec, 72 ºC for 60 sec, and extension at 72 ºC for 10 min. The cycling conditions were chosen so that none of the templates analyzed reached a plateau at the end of the thermal cycling, i.e., they were in the exponential phase of amplification. PCR products were separated on a 2% agarose gel and photographed using UV illumination to visualize ethidium bromide staining. Images were inverted in Adobe Photoshop.

7.2. List of PCR primers used in this study

Gene	Forward primer (5' ->3')	Reverse primer (5' ->3')
GAPDH	TTCTTTTGCGTCGCCAGCCGA	GTGACCAGGCGCCCAATACGA
HDAC2	TGGCCGGGGAGCCCATGG	GGCAGTGGCTTTATGGGGCCTATAT
TF	CGGAAGATGAGGCTCGCCGT	TGCTCCGACACTGCACACCA
RGS16	TGGCAAGTTCGAGTGGGGCA	GGAAAGCGTGGAAGGCAGCCA
GABRA6	TTTGGGCCCGTGTCAGATGTGG	GGTGTCAGGCGTCCAGATTTTACTG
RGS9	GGACTACGGCCTGGACCGAGT	ACAATCCCTCCCAGGGACACAGA
PCDH11Y	GGCCTGCCCCTTGGCTATCC	GGCCTCTTCCACAGTTGGTTGAACA
ANKRD32a	TGTTTGTTGCAGAGGCAGTC	GACCCAAGCACACCAGTTTT
ANKRD32b	ACTAGGGGAGACAATGCCTGT	CCCACCAATCCATACTGCTT
NLGN4X (exon 7)	TGAGACTCACAGGCGCCCCA	CGTGTGCCTGCAGCGACTCA
NLGN4X (3'UTR)	CCCTCTGCCCTACCCGCTCA	GCCGGGATGGGATGACTGCC
GLIPR1L2	CCTGGCAGTAGGGGGCGTTT	TCCCAAGTCATGAAGCGCAAGTTAG
Rpl32	CAACATCGGTTATGGGAGCAACA	TGACATTGTGGACCAGGAACTT
Ankrd32a	CAGGGCCCCACTAGTACGTGC	AGCTCTGGCCGACCTTCTGCT
Ankrd32b	GGCATGGGTCCAGACCTTCTAGTT	GCGTCCAGCCAGCATTGTCTT
IGF2	Hs00171254_m1 (Applied Biosystems)	
GAPDH	Hs99999905_m1 (Applied Biosystems)	

7.3. Immunohistochemistry and histochemistry

For neuropathological evaluation and validation studies, brain tissue samples were fixed in 4% PFA for 2–3 days at 4 °C. Following fixation, tissue was cryoprotected in graded sucrose solutions (up to 30%) at 4 °C, then frozen at -40 °C in 2-methylbutane (J.T. Baker), and stored at -80 °C. Alternatively, fixed tissue samples were paraffin-embedded for routine neuropathological evaluations.

Frozen tissue samples were cut at 60 µm using a Leica CM3050S cryostat and either mounted onto Superfrost/Plus slides (Fisher Scientific Co.) or used free-floating. Tissue sections were incubated in 1% hydrogen peroxide/PBS to quench endogenous peroxidase activity. Sections were washed in PBS (3 x 15 min) and incubated in blocking solution containing 5% (v/v) normal donkey serum (Jackson ImmunoResearch Laboratories), 1% (w/v) bovine serum albumin, and 0.4% (v/v) Triton X-100 in PBS for 1 h at room temperature. Primary antibodies were diluted in blocking solution and incubated with tissues sections overnight at 4 °C. Sections were washed with PBS (3 x 15 min) prior to being incubated with the appropriate biotinylated secondary antibodies (Jackson ImmunoResearch Labs) for 1.5 h at room temperature. All secondary antibodies were raised in donkey and diluted at 1:200 in blocking solution. Sections were subsequently washed in PBS and incubated with avidin-biotin-peroxidase complex (Vectastain ABC Elite kit; Vector Laboratories) for 1 h at room temperature. Finally, sections were washed in PBS (3 x 15 min) and signals were developed using a DAB peroxidase substrate kit according to the manufacturer's protocol (Vector Laboratories). Following washes in PBS, sections were mounted on Superfrost Plus charged slides, dried, dehydrated, and cover slipped with Permount (Fisher Scientific Co.). Sections were digitized using ScanScope scanner (Aperio). Digitized images were assembled in Adobe Photoshop and Illustrator.

NADPH-diaphorase histochemistry was performed as previously described[69].

8. Dissection of the Mouse Neocortex

We performed qRT-PCR in orthologous areas of the mouse NCX at equivalent developmental periods using a mathematical model to translate neurodevelopmental time across mammalian species[70] (www.translatingtime.net). For analysis of mouse *Ankrd32* by qRT-PCR, CD1 mice were obtained from Charles River Laboratories. Mice were sacrificed according to institutional regulations. Brains from embryonic day (E)15, postnatal day (P) 3 and adult mice (3 brains per period) were dissected either fresh in cold PBS or after storage of the whole brain in RNA*later* reagent (Qiagen). Dissected regions were sampled according to the "Prenatal Mouse Brain Atlas" (for E15 mouse brain) and "Atlas of the Developing Mouse brain" (for P3 and adult mouse brain)[71,72]. The following regions were dissected:

E15:

MFC – sampled from medial part of the frontal cortex.

DLFC – sampled from dorsolateral part of frontal cortex (containing prospective primary motor cortex).

S1C – sampled from anterior part of parietal cortex.

A1C – sampled from prospective primary auditory cortex (upper part of prospective temporal cortex).

V1C – sampled from occipital part of the hemisphere.

P3 and **adult**:

OFC – sampled from orbital part of the frontal cortex after removal of olfactory bulb (containing VO and LO areas according to Paxinos et al.).

MFC – sampled from medial part of the frontal cortex (containing PrL and IL areas according to Paxinos et al.).

M1C – sampled form the dorsal part of the frontal cortex (containing M1 and M2 areas according to the Paxinos at al.).

S1C – sampled from the rostral part of the parietal cortex (containing all S1 area subdivisions according to Paxinos et al.).

A1C – sampled from the upper part of the temporal cortex (containing AuD, Au1 and AuV areas according to Paxinos et al.).

V1C – sampled from the occipital part of the hemisphere (containing V1, V2ML, and V2L areas according to Paxinos et al.).

RNA extraction of P3 and adult samples was performed following the same protocol used for human tissue samples (see Supplementary Information 5.1). Due to the small amount of tissue for E15 samples, a Trizol (Invitrogen) extraction protocol was performed according to manufacturer's instructions. Eighty ng of total RNA was used for cDNA synthesis, and quantitative RT-PCR was performed and analyzed as the same method with human brain samples. Ct values were normalized to the housekeeping gene *Rpl32*. Three technical replicates were performed on each sample.

9. Expression Quantitative Trait Loci Analysis

The genotyping results consisted of 2.5 million SNPs for each of 57 brains using Illumina Omni-2.5 SNP arrays. We used PLINK to calculate the minor allele frequency (MAF), Hardy-Weinberg p-value and SNP-wise call rate for each SNP. The eQTL analysis was restricted to 890,000 SNPs that had MAF > 0.1, Hardy-Weinberg p-value > 0.05, and

SNP-wise call rate greater than 0.99 in order to reduce genotyping errors and increase statistical power. The analysis was further restricted to SNPs within 10 kb of the transcription start or stop sites of the 15,132 expressed genes.

The expression data in 11 NCX areas were averaged to get a single expression value in NCX for each gene, and *cis*-eQTL analysis was conducted separately for each region as described below. We adjusted for other non-genetic factors impacting gene expression measures in two ways. First, our phenotype was the expression level relative to the average developmental trend for each gene (see below); this allowed us to focus on the individual variation independently of the developmental trends. Second, we included possible confounding factors as covariates in our linear model for relative expression; these covariates included RIN, PMI, sex, the first two principal components of the global transcriptome, and ethnicity.

In order to compare individual differences in single gene expression across developmental time points, we adjusted the expression values as follows. For each gene, we sorted the expression values in 57 brains according to their age and fitted a smooth function (computed by local regression or LOWESS) to obtain the average expression levels over development stages. We calculated a relative expression level by subtracting the value of this trend curve from the normalized individual gene expression value.

After separating genes and SNPs by chromosome, PLINK was used to conduct a quick scanning for possible associations between all genes and SNPs near the beginning and end of the gene as described above with the following arguments: 'plink --file myfile --pheno expression.txt --all-pheno --assoc --pfilter 1e-3 '. Using these phenotypes, we ran linear regression using PLINK with following settings: “plink --file selectfile --pheno select_expression.tx --sex --all-pheno --covar covariants.csv --covar-name PC1, PC2, PMI, RIN, Cauc, Afric, Asia, Hisp --linear --pfilter 1e-3”.

A conservative two-level strategy was applied for multi-test correction. We performed a Bonferroni correction for all tests involving the same gene to obtain a gene-wise adjusted p-value for the SNP with the lowest p-value. We estimated the genome-wide q-value (FDR) for all 15,132 genes according to the Benjamini-Hochberg formulation, employing the R function p.adjust. We selected those genes and SNPs with q-value less than 0.1 for our final list of detected *cis*-eQTLs.

In order to check the distribution of SNP distance from either a RefSeq annotated transcription start site (TSS) or transcription end site (TES) of the associated gene for all SNPs that passed gene-wide correction without genome-wide correction, we tested all SNPs within 1Mb of TSS and TES and then used Bonferroni correction to obtain the gene-wide adjusted p-value. Those SNPs with gene-wide p-value <0.05 were collected and their distances from TSS and TES were calculated.

10. Supplementary Tables

Supplementary Tables are provided in a single Microsoft Excel file on the CD rom.

Supplementary Table 1. Donor/specimen metadata

This table summarizes donor information, including sex, age, ethnicity, pH, PMI, cause of death, and medical history for each human brain subject. Brain tissue was obtained from clinically normal donors and multiple brain specimens per period were analyzed to reduce the effects of individual variation. Demographic and quality control characteristics were as follows: age, 5.7 post-conceptual weeks (PCW) to 82 years (Y); sex, 31 males and 26 females; PMI, 12.11±8.63 (mean±SD) hours; pH, 6.45±0.34 (mean±SD).

Supplementary Table 2. List of CNVs per specimen

Complete list of predicted copy number variations (CNVs) in each individual donor. Putative CNVs were predicted by two different algorithms (cnvPartition in GenomuStudio and PennCNV) using SNP intensity data generated by Illumina Human Omni 2.5-Quad Bead Chips. After validation of genomic normality, 57 brains were used in the transcriptome analysis.

Supplementary Table 3. Tissue sample metadata

This table summarizes quality control data for individual tissue samples, including RNA integrity number (RIN) and tissue dissection score (DS). Averaged RIN for each sample was 8.83±0.93 (mean±SD), and only those with a RIN above 5 and DS above 2 were processed for exon array analysis.

Supplementary Table 4. List of exon array CEL files

List of all exon array CEL files that passed the quality control steps throughout microarray cDNA preparation, hybridization, and outlier detection (Supplementary Information 5). A total of 1,340 samples were considered in downstream analyses.

Supplementary Table 5. List of region-enriched DEX genes

This table shows spatially DEX genes that were enriched in or restricted to only one of the regions (region-enriched genes). Using a combination of a sliding window algorithm and two-sample t-test, region-enriched DEX genes were analyzed within the periods of 3-7, 4-8, 5-9, 6-10, 7-11, 8-12, 9-13, 10-14, 11-15, respectively. The fold differences were calculated by comparing the gene expression value from one region of interest with another region that showed the second highest expression value. A gene with greater than 2-fold differences in a given window was considered to be region-enriched. The post

hoc group assignment analysis was performed by calculating the fold differences attributed to the two periods bordering each side of the developmental groups (6-7 and 8-9 for prenatal-postnatal group; 11-12 and 13-14 for postnatal adult group). If the fold difference was greater than 1.5, the gene was assigned to that developmental group, resulting in either one or both group(s) being assigned.

Supplementary Table 6. Tukey's pairwise analysis of human and mouse *ANKRD32* isoforms

The Tukey's pairwise comparison was performed after ANOVA to determine which pairs of NCX areas were significantly different from each other. The TukeyHSD function in R was used to calculate the gene expression mean difference and corresponding p-value.

Supplementary Table 7. List of sex-biased DEX genes

Differentially expressed genes between males and females were analyzed by a combination of a sliding window algorithm and t-test. The window size was set as 5 in order to ensure adequate representation of both sexes, and periods 2-6, 3-7, 4-8, 5-9, 6-10, 7-11, 8-12, 9-13, 10-14, 11-15 were examined. For each window, a two-sample t-test was applied to determine the sex-biased genes in the same region (cutoff of FDR 0.01 and 2-fold difference). The post hoc group assignment analysis was performed by calculating the fold differences attributed to the two periods bordering each side of the developmental groups (6-7 and 8-9 for prenatal-postnatal group; 11-12 and 13-14 for postnatal-adult group). If the fold difference was greater than 1.5, the gene was assigned to that developmental group, resulting in either one or both group(s) being assigned the gene.

Supplementary Table 8. List of sex-biased DEU genes

Genes with differential exon usage between males and females were analyzed by a combination of splicing index and two-sample t-test under a sliding window algorithm. The window size was set as 5 in order to ensure adequate representation of both sexes, and periods 2-6, 3-7, 4-8, 5-9, 6-10, 7-11, 8-12, 9-13, 10-14, 11-15 were examined. For each window, exon cluster expression was normalized by dividing its gene-level expression. The splicing index was then calculated as a "fold change" between the normalized levels of each exon cluster to determine the sex-specific DEU (cutoff of FDR < 0.01 and 2-fold difference).

Supplementary Table 9. Co-expression network modules

List of 29 modules corresponding to spatiotemporal patterns identified by weighted gene co-expression network analysis (WGCNA) with its functional interpretation, list of hub genes, and result of gene ontology analysis performed by DAVID algorithm. Top 10 hub genes are listed for each module, though some consist of less than 10 due to small module sizes. Two large modules (M20; 3,287 genes and M2; 2,745 genes) appeared to be simultaneously co-regulated across all regions with opposite developmental trajectories (Supplementary Figs 21 and 22). In addition to the predominant temporal patterns, region-specific modules were also identified, such as M9 (MD), M19 (CBC), M23 (STR). Several modules exhibited no obvious spatiotemporal expression patterns but were enriched in genes associated with cell cycle (M1), cytoskeleton (M3), olfactory receptors (M13), and sex-specific transcription (M22).

Supplementary Table 10. Correlation between network modules and neurobiological categories

The correlation coefficient was calculated between each network module and functional gene group of specific neurobiological categories. Manually generated annotations for various functional groups and consisting gene list are also provided in Supplementary Table 12. The first principle component of each module expression (eigengene by WGCNA analysis) and each gene group were applied for correlation coefficient computation. Notably, M8 module (high in prenatal NCX, HIP, and AMY) showed strong correlation with layer 2 to 6 cortical glutamatergic neurons, while M15 module (high in postnatal NCX, HIP, AMY, and STR) showed strong correlation with dendrite development, synapse development, glial cells (astrocytes, myelination, and oligodendrocytes), GABA shift, and certain cortical GABA cell types. Also, M20 module (high in prenatal) showed strong correlation with cell proliferation, neuron differentiation, neuron migration, and cortical deep layer glutamatergic neurons, while it shows strong anti-correlation with synapse development, dendrite development, glial cells (astrocytes, microglia, oligodendrocytes, and myelination), GABA shift, and certain cortical GABA cell types. On the contrary, correlation of M2 module (high in postnatal) showed opposite trend with these functional groups.

Supplementary Table 11. Correlation between network modules and confounders

This table shows correlation between network modules and all the factors associated with the samples, including possible confounders (RIN, PMI, pH, and DS). To check the effects of RIN, PMI, pH, and dissection on module identification, the coefficient of determination (r^2) and *P* value was calculated after constructing a linear ANOVA model. The modules

with r^2 of any factor greater than 0.1 and r^2 of any confounder (RIN, PMI, pH, and DS) larger than the largest r^2 of the main factors (age, region and sex) were considered as modules corresponding to confounders. Five modules were detected as confounder-related modules and removed from further analyses. One module (grey), consisting of genes not assigned to any other module, was also removed.

Supplementary Table 12. Gene lists for neurobiological categories

List of the manually generated gene group annotations for various neurobiological categories. We summarized the gene list into three major categories: neurodevelopmental processes, neural cell types, and major neurotransmitter systems. Each category was classified into functional groups and subgroups based on existing GO categories (www.geneontology.org) and published data from model organisms and human medical genetics.

Supplementary Table 13. Genes correlated with genes involved in neurodevelopmental processes

The correlation coefficient was calculated between each functional gene group involved in specific neurodevelopmental processes and all genes expressed. The first principle component of each gene group was computed against each gene expression level. Top 100 genes are listed.

Supplementary Table 14. Genes correlated with genes expressed in specific neural cell types

The correlation coefficient was calculated between each functional gene group involved in specific neural cell types and all genes expressed. The first principle component of each gene group was computed against each gene expression level. Top 100 genes are listed.

Supplementary Table 15. Transcripts correlated with ASD-associated genes

This table shows the top 50 correlated transcripts with autism spectrum disorder (ASD) related genes, which were compiled from recent reviews and relevant databases (Supplementary information 6.10). Correlation coefficient (r^2) was calculated between the expression level of each gene and all transcripts expressed. Functional annotation suggested that these genes were predominantly associated with phosphoprotein (Bonferroni-adjusted $P=1.3 \times 10^{-30}$), synapse (Bonferroni-adjusted $P=9.1 \times 10^{-30}$), synaptic transmission (Bonferroni-adjusted $P=5.8\times 10^{-26}$), cell junction (Bonferroni-adjusted $P=1.5 \times$

10^{-20}), neuron projection (Bonferroni-adjusted P=3.4 × 10^{-19}), and ionic channels (Bonferroni-adjusted P=3.8 × 10^{-17}).

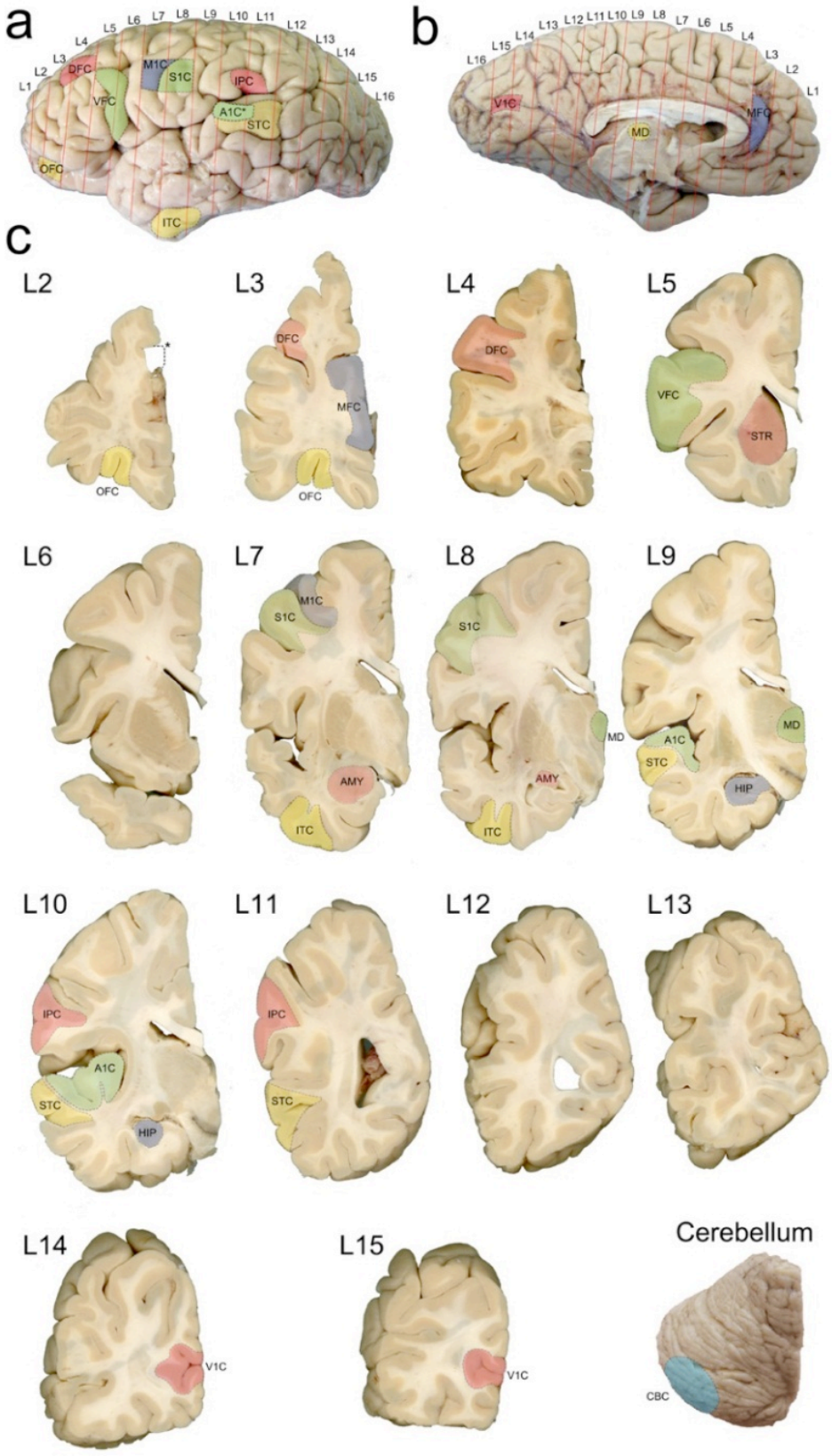
Supplementary Table 16. Transcripts correlated with schizophrenia-associated genes

This table shows the top 50 correlated transcripts with schizophrenia related genes, which were compiled from recent meta-analyses and relevant databases (Supplementary information 6.10). Correlation coefficient (r^2) was calculated between expression level of each gene and all transcripts expressed. Functional annotation suggested that these genes were mainly associated with glycoprotein (Bonferroni-adjusted P=1.5 × 10^{-16}), plasma membrane (Bonferroni-adjusted P = 1.5 × 10^{-12}), ionic channel (Bonferroni-adjusted P=2.6 × 10^{-9}), synaptic transmission (Bonferroni-adjusted P=3.8 × 10^{-7}), and phosophoprotein (Bonferroni-adjusted P=5.3 × 10^{-7}).

Supplementary Table 17. Regional *cis*-eQTLs

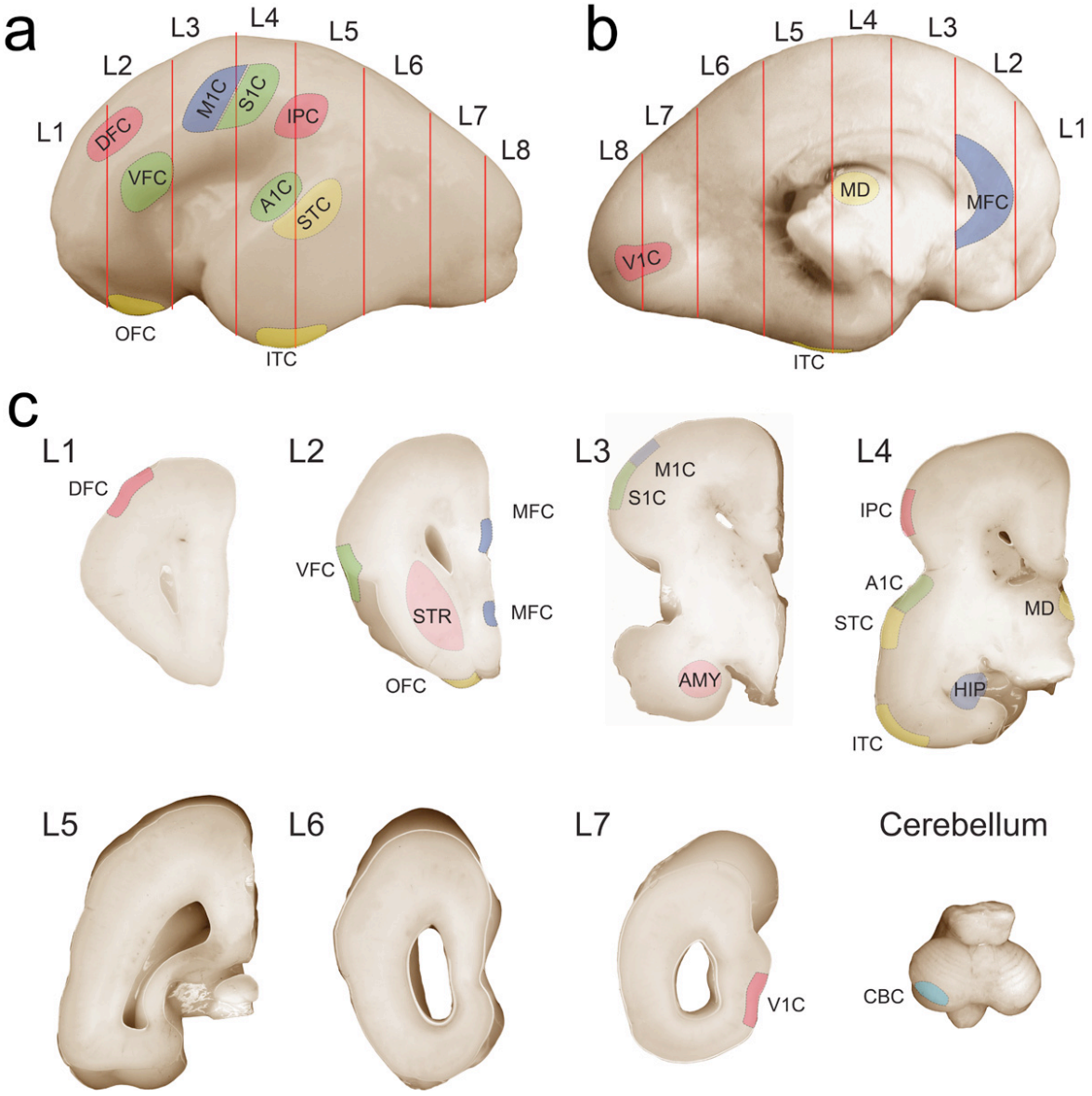
This table shows *cis*-eQTL within 10kb flanking two ends of the target gene, detected in 6 brain regions (NCX, HIP, AMY, STR, MD and CBC). Rows in red are eQTLs with genome-wide FDR < 0.1, rows in black are eQTLs with genome-wide FDR < 0.2. SNP count is the number of local SNPs tested for the gene. MAF is minor allele frequency in our dataset. HWE p-value is the Hardy-Weinberg-Equilibrium test p-value of our genotyping result. Number of BB calls is the number of individuals with genotype of homozygous minor allele. WALD p-value is the original association p-value calculated by PLINK. Covariants p-value is the WALD p-value with co-factors of RIN, PMI, sex, ethnicity, and first two principle components. Gene-wide p-value was calculated from the original WALD p-value after Bonferroni correction by number of local SNPs. Permutation p-value was calculated from label-swapping adaptive permutation by PLINK with maximum of 1,000,000 times. FDR p-value was calculated from the Gene-wide p-value after genome-wide FDR correction.

11. Supplementary Figures



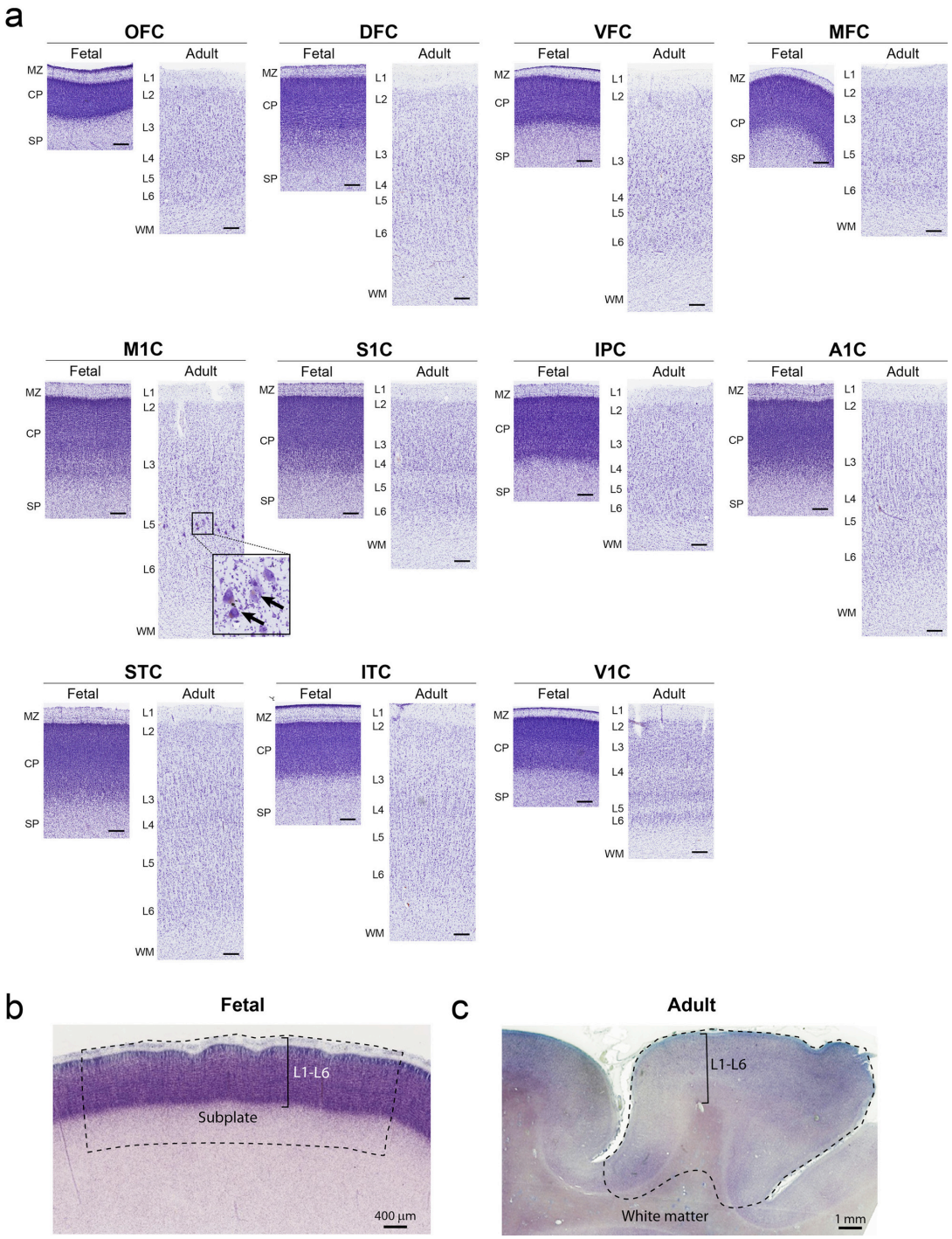
Supplementary Figure 1 | Demarcation of adult brain regions and NCX areas. a-c, Representative adult human brain images from period 14 are shown illustrating the

sampling locations depicted on the lateral (a) and medial (b) surfaces of the hemisphere and the dorsal surface of the cerebellar hemisphere (c). Relative sizes of the sampled regions of interest are depicted on coronal slices (c, L2 – L15) of the left cerebral hemisphere. Red lines in (a) and (b) represent locations of coronal sections. The posterior side of the slice is shown, and the average section thickness is 1 cm. Brain regions and NCX areas of interest are represented by different colors and two or three letter abbreviations (for full names see section 2.2 of the Supplementary Information). A1C* is located on the temporal bank of the lateral sulcus (i.e., planum temporale and transverse gyrus of Heschl), and cannot be observed on the lateral view of the hemisphere; for illustrative purposes only, relative size and position of the A1C* was depicted on the lateral surface. The dotted line and asterisk (L2) depict an artifact during tissue processing. For detailed sampling procedures see section 2.2 of the accompanying Supplementary Information. These images were generated using fixed brain specimens not used to generate transcriptome data.



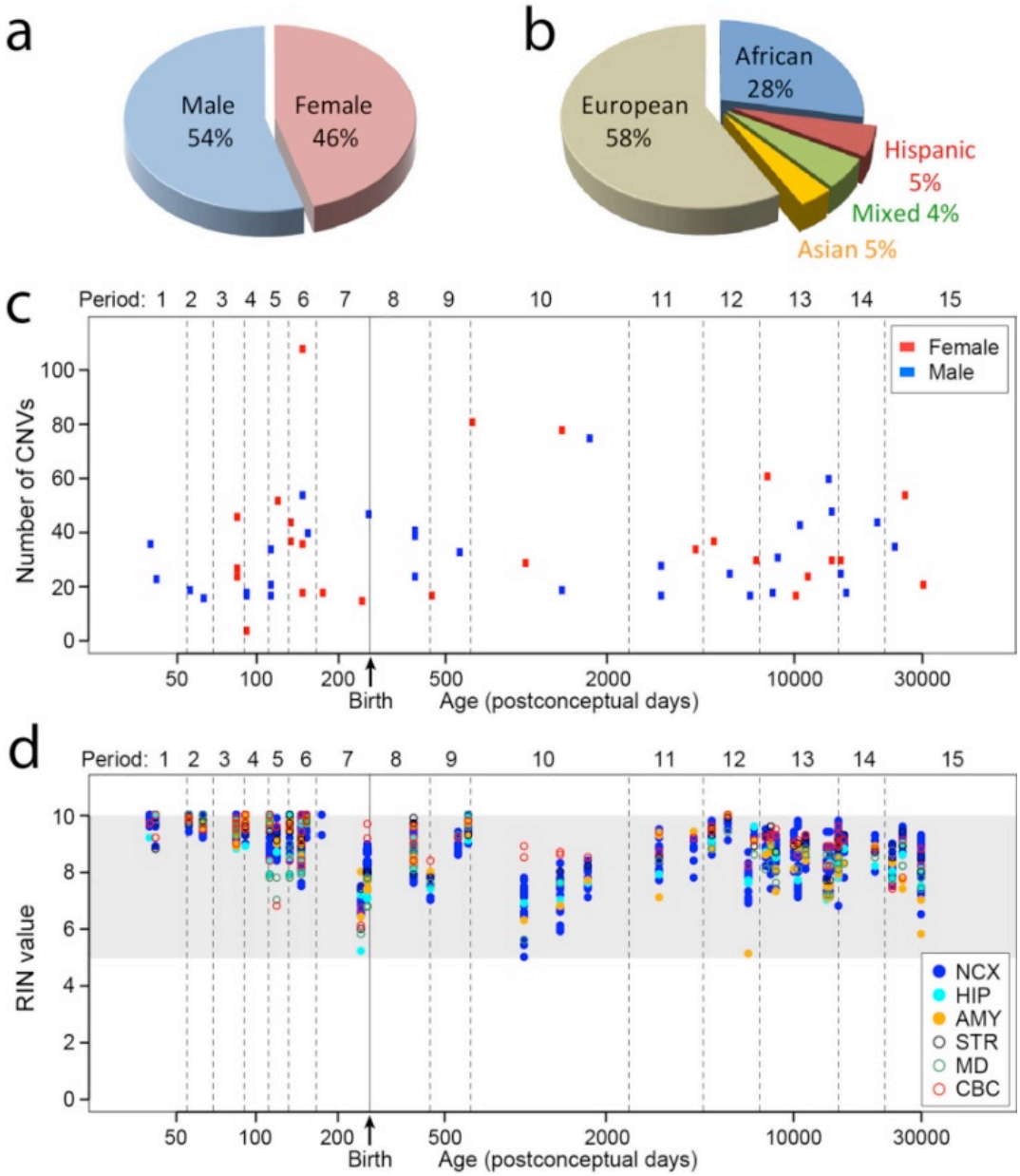
Supplementary Figure 2 | Demarcation of fetal brain regions and NCX areas. a-c, Representative fetal human brain images from period 6 are shown illustrating the sampling locations of tissues used for transcriptome analysis. Relative positions of the regions of interest are depicted on the lateral (a) and medial (b) surfaces of the hemisphere and the dorsal surface of the cerebellar hemisphere (c). Relative sizes of the sampled regions of interest are depicted on coronal slices (c, L1 – L7) of the left cerebral hemisphere. Red lines in (a) and (b) represent locations of coronal sections. The posterior side of the slice is shown, and the average section thickness is 0.5 cm. Regions of interest are represented by different colors and two or three letter abbreviations (for full names see section 2.2 of the Supplementary Information). In total, 16 regions of interest were sampled as follows: 5 regions from the frontal lobe cortex (OFC, DFC, VFC, MFC

and M1C), 2 regions from the parietal lobe cortex (S1C and IPC), 4 regions from the temporal lobe cortex (A1C, STC, ITC and HIP), 1 region from the occipital lobe cortex (V1C), 3 regions from subcortical structures (MD, AMY and STR) and 1 region from the cerebellum (CBC). Sampled regions of interest always contained cortical plate and part of the underlying subplate zone. For detailed sampling procedures see section 2.2 of the Supplementary Information. These images were generated using fixed brain specimens not used to generate transcriptome data.

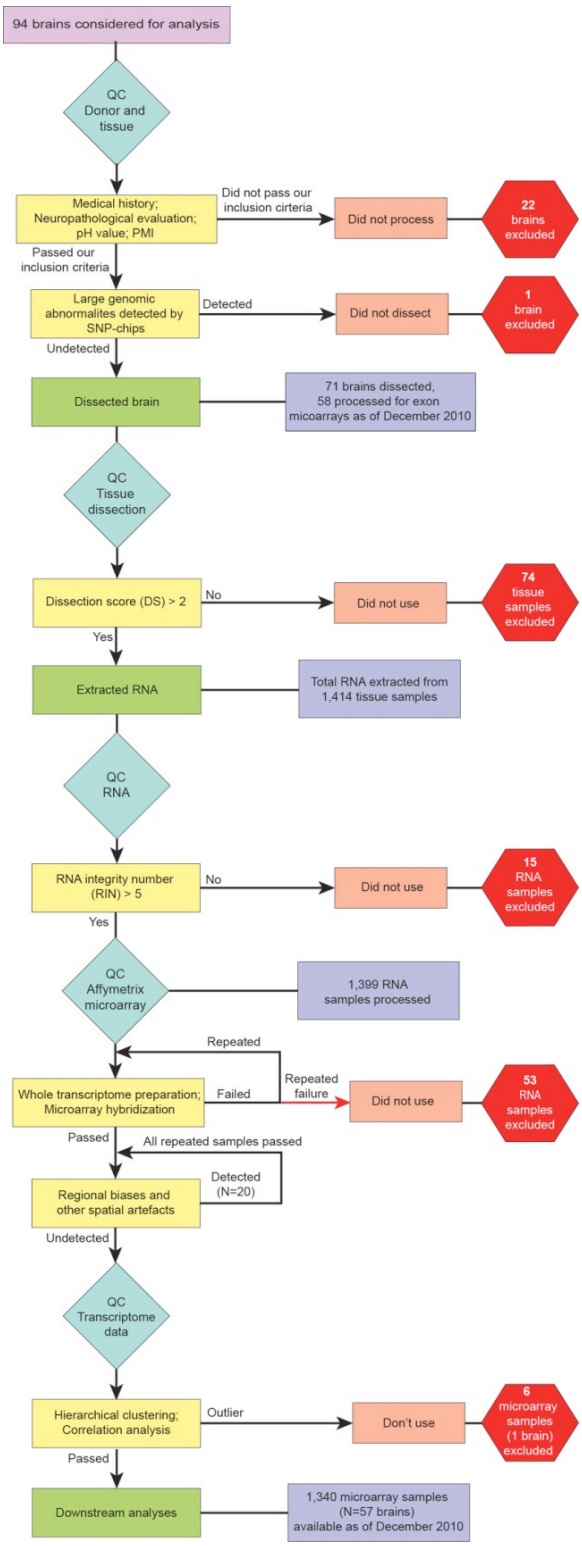


Supplementary Figure 3 | Histological evaluation and dissection of NCX areas. a, Nissl staining was used to validate the following NCX areas: OFC, DFC, VFC, MFC, M1C, S1C, IPC, A1C, STC, ITC, and V1C. Sampling from fetal (period 6) and adult (period 14) specimens was carried out according to the methods and criteria used for the dissection of exon array samples. Staining in fetal NCX areas revealed densely packed neurons in the cortical plate (CP) with no obvious distinctions among cortical layers. The marginal zone (MZ) and subplate (SP) were present in the fetal samples. Conversely, adult NCX neurons

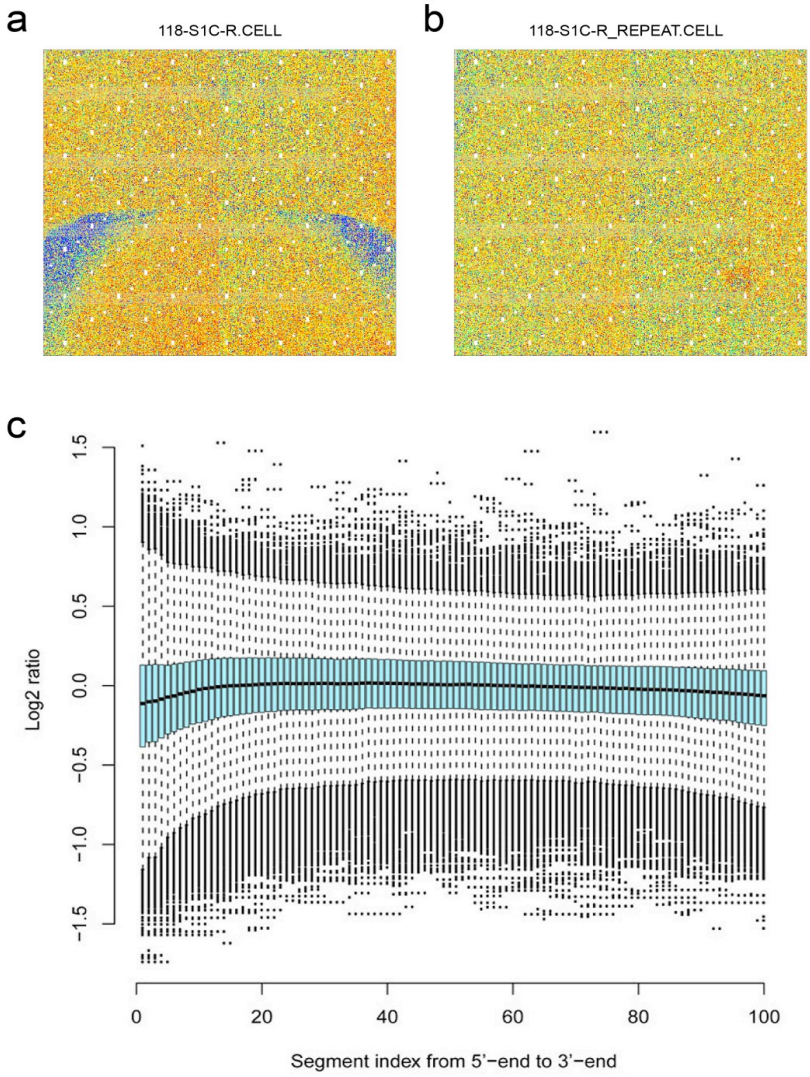
were at a lower density and were arranged in layers (L1–L6) containing neurons of distinct size and shape (e.g., L5-specific Betz cells in M1C; arrow). The adult cortical layers had a greater thickness than the fetal cortical plate, and white matter (WM) was apparent below the laminar pattern. Scale bars, 250 μm. **b, c,** Nissl staining images of the fetal (b, period 6) and adult (c, period 14) VFC are shown to illustrate the microscopic boundaries of the NCX tissue sampled for transcriptome analysis. In the fetal brains (b), the entire cortical plate (L1–L6) and adjacent subplate zone was sampled. In the adult brains (c) all six cortical layers (L1–L6) and underlying gyral white matter were sampled. Dotted lines in (b) and (c) represent dissection boundaries. These three images were generated using fixed brain specimens not used for the exon array analysis.



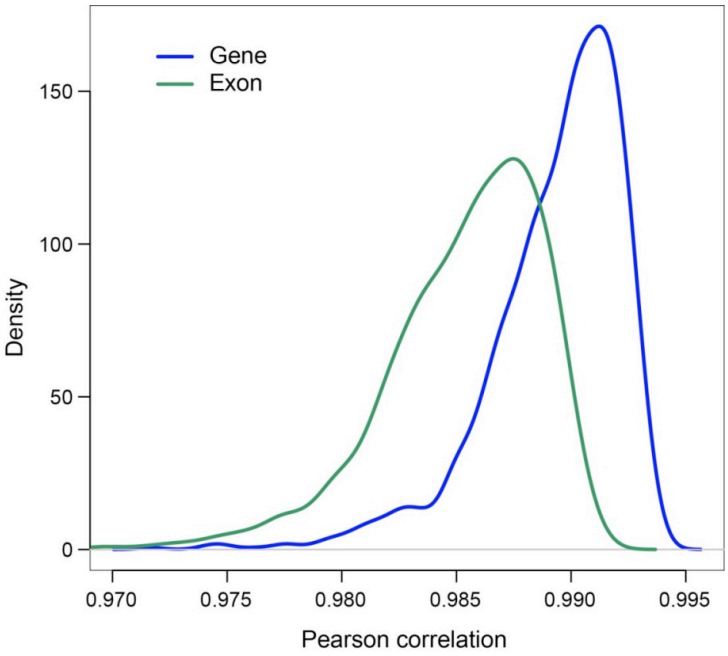
Supplmentary Figure 4 | Demographics, genotyping, and RNA integrity of analyzed brain specimens. **a**, Sex distribution of specimens. **b**, Ethnicity distribution of specimens according to SNP genotypes. **c**, Number of CNVs per individual specimen distributed across sexes and time periods. **d**, RNA quality was evaluated by determining the RNA integrity number (RIN) for each sample (8.83±0.93, mean±SD), and only those with a RIN above 5 were processed (1,340 samples). Age is represented as $\log_2$–transformation of postconceptional days. Time of birth, separating prenatal from postnatal periods, is marked by vertical solid line.



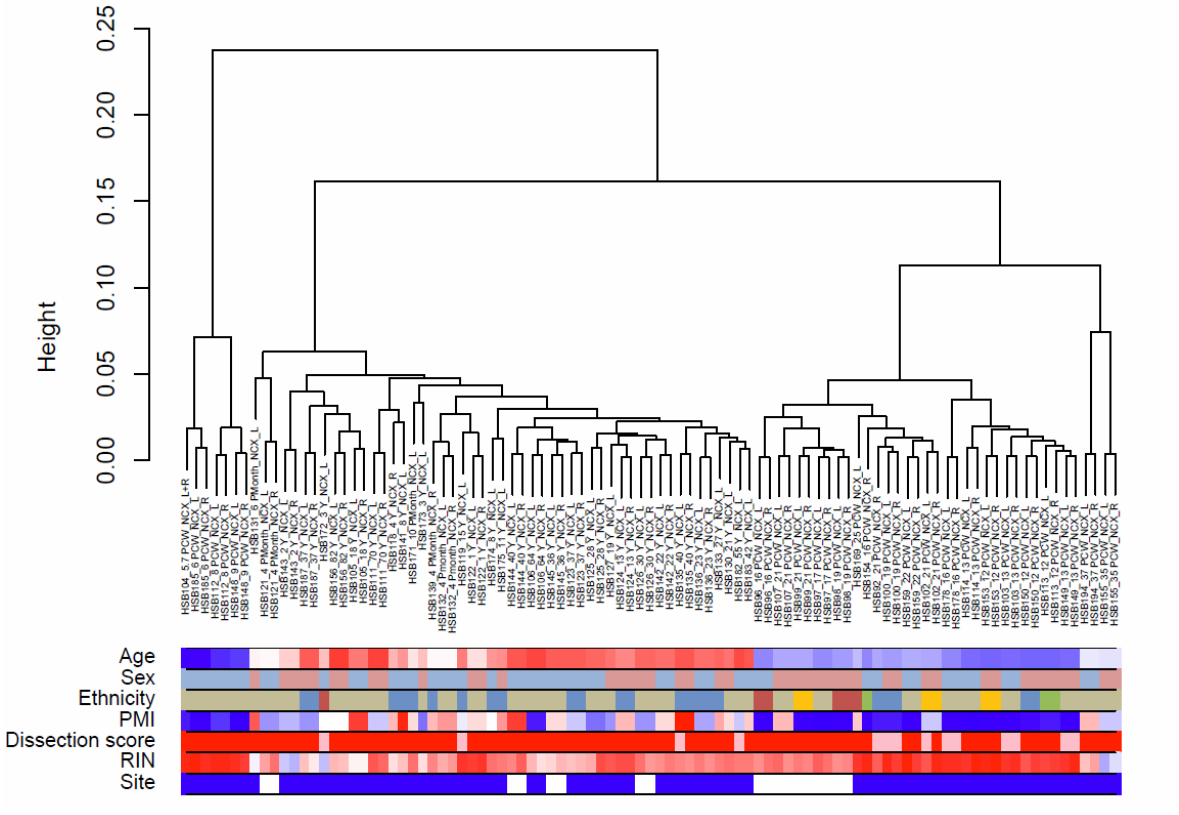
Supplementary Figure 5 | Quality control workflow. Blue diamonds represent a quality control (QC) step, yellow squares represent criteria used to determine quality of the samples, green squares represent actions taken after quality control, red hexagons represent samples excluded from further analysis, and purple squares represent samples considered for further analysis. For details see Supplementary Information 2.2.



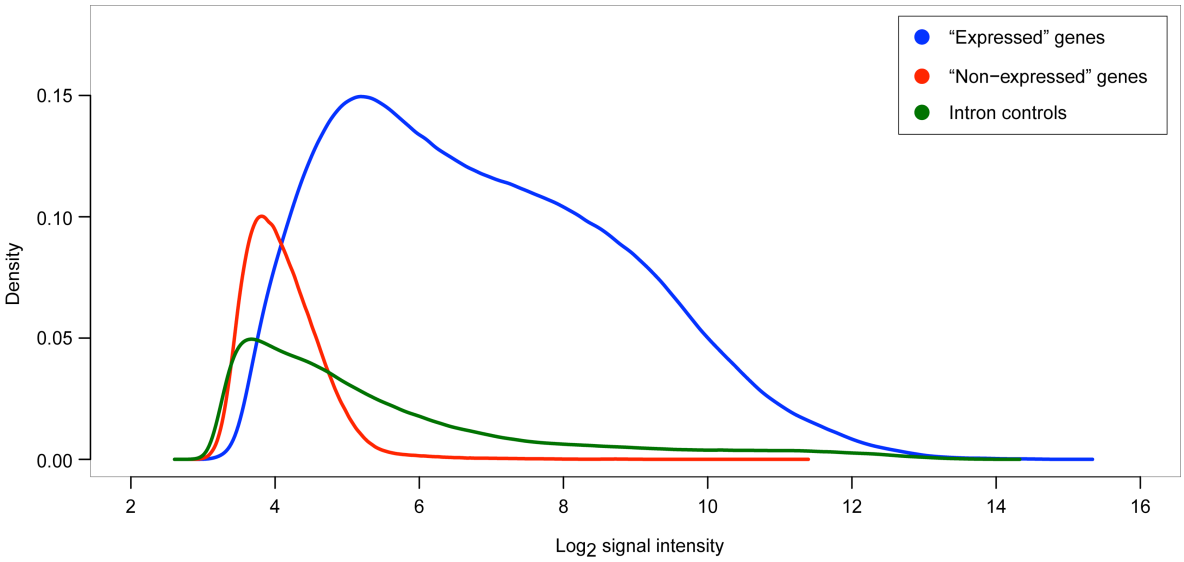
Supplementary Figure 6 | Exon array hybridization quality control. a, b, The color of each dot in the plot represents relative intensity of a probe in the target chip compared to the average intensity of that probe in the same batch of chips. Individual probes of one gene were randomly located on the whole chip, and consequently there should be no obvious spatial pattern in the plot. The relative intensity plot of a low quality chip for sample 118-S1C-R with spatial artifacts (a) and the relative intensity plot of the high quality replicate chip for the same sample (118-S1C-R_REPEAT), without any obvious spatial artifact (b). **c,** The longest transcript of each gene was split into 100 segments with equal length, from 5'-end to 3'-end, and the log2-transformed signal intensity ratio for each segment was calculated compared with the expression of the whole gene across 1,340 chips. The statistics of the ratios for all genes were displayed by a box plot for each segment. The flat solid line consisting of the mean lines of the box plots displays the array hybridization uniformity.



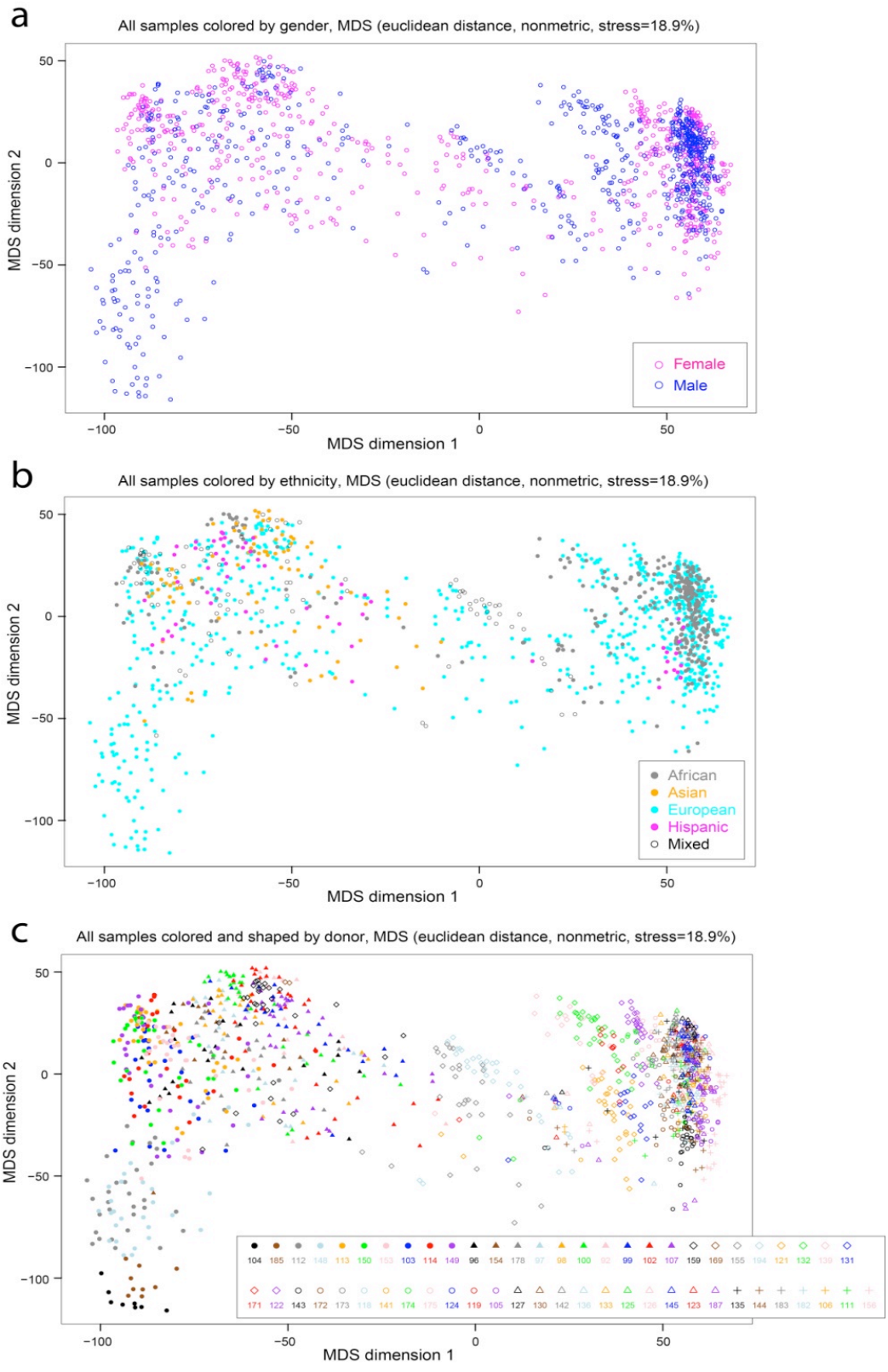
Supplementary Figure 7 | Expression analysis with and without SNP-containing probe sets. The gene and exon expression were evaluated twice, including and excluding 105,271 probes with SNPs. The Pearson correlation analysis was performed to show the degree to which SNPs affect the expression data. The density plots give the distribution of Pearson Correlation Coefficient (PCC) values for all 1,340 samples (gene and exons are represented by blue and green lines, respectively). The majority of PCC values (98.8%) were within the range of 0.980-0.995 for gene and of 0.975-0.995 for exon, suggesting that SNPs have a minimal effect on gene and exon expression.



Supplementary Figure 8 | Hierarchical clustering of NCX samples. Dendrogram branching shows hierarchical sample clustering of NCX. The colors underneath label the potential confounders: age (blue to red representing younger to old), sex (light blue, male; pink, female), ethnicity (African, blue; African/European, green; Asian, yellow; European, light brown; Hispanic, dark red), PMI (low to high representing blue to red), dissection score (low to high representing blue to red), RIN (low to high representing blue to red), and the array processing site (Site 1, blue; Site 2, white). The samples clustered predominately by age (periods) and not by sex, ethnicity, PMI, dissection score, RIN, or the array processing site. Three major clusters (a stage 1 and 2 cluster, a prenatal cluster, and a postnatal cluster) were found in the NCX. Clustering data for other regional samples (HIP, AMY, STR, MD and CBC) are not shown but revealed the same trend.

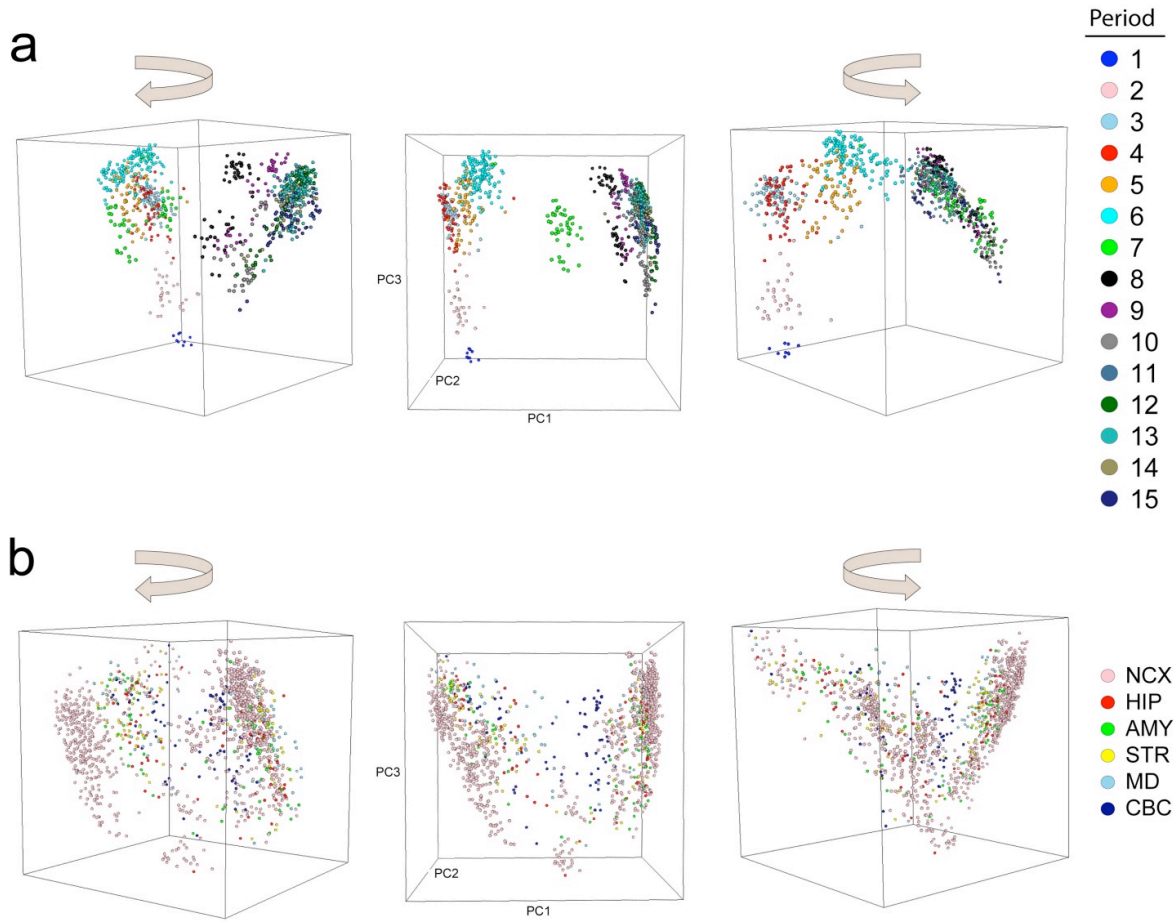


Supplementary Figure 9 | Distribution plot of exon array signal intensities for "expressed", "non-expressed", and intronic controls. The "expressed" genes are defined by two criteria of log_2-transformed signal intensity and DABG p-value (Supplementary information 6.1). The first group of genes included genes with averaged DABG P <0.01 in at least one region and log2 intensity ≥6 in at least one sample. These genes were considered as "expressed" genes (blue). The second group included genes whose averaged DABG P >0.01 or log_2 intensity <6 in all samples and were considered "non-expressed" group (red). The final group is intronic controls (dark green) consisting of probe sets designed from intron regions of a set of housekeeping genes. Note the distinct distribution of "expressed" genes compared to "non-expressed" genes and intronic controls.

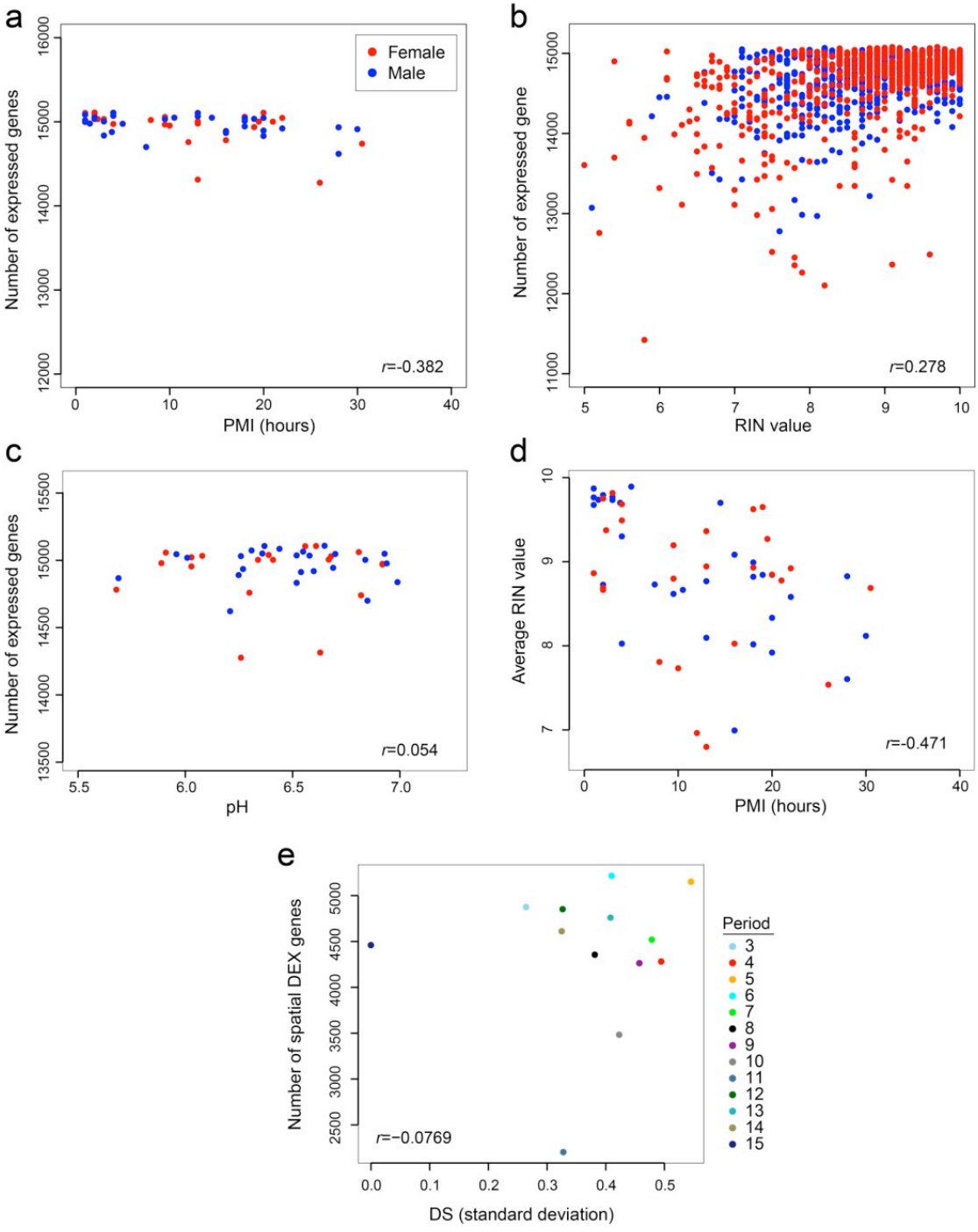


Supplementary Figure 10 | Multi-dimensional scaling according to sex, ethnicity, and donor. Two-dimensional multi-dimensional scaling plot showing genome-wide transcriptional similarity between any two samples. Each sample is represented as a single point. Proximity indicates transcriptional similarity between two samples. Euclidean

distance of log_2-transformed signal intensity (expression) values was used to measure the pairwise dissimilarity. The isoMDS function in the R package was used to create the configuration of all points in the two dimensional space. **a,** colors indicate sex. **b,** colors indicate ethnicity. **c,** colors and shapes indicate donor.

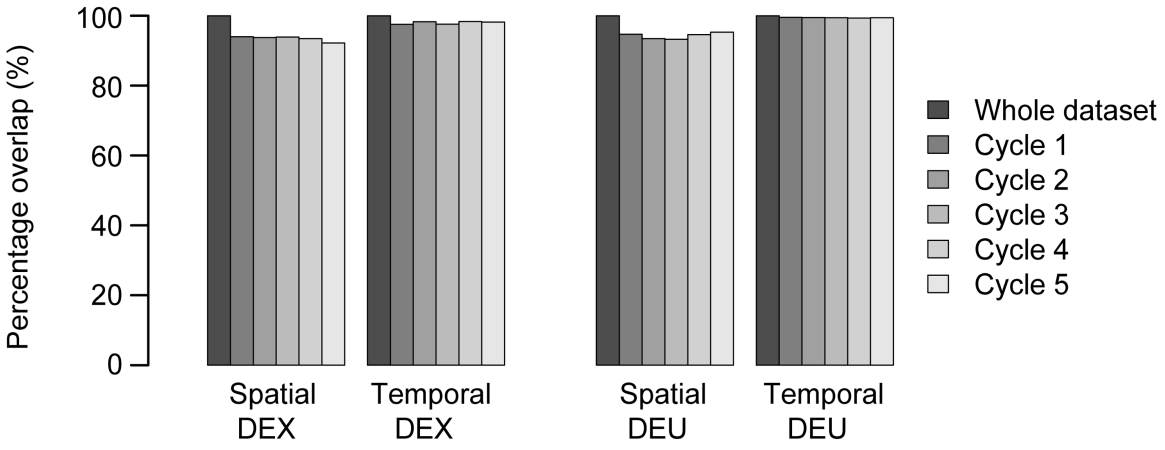


Supplementary Figure 11 | Principal component analysis. a, Three-dimensional plot of PCA of NCX samples across 15 time periods. Each point represents one NCX sample. Samples are colored by period. Three dimensional figures are rotated to provide a better view of the separation within prenatal (left) and postnatal (right) periods. **b**, PCA plot for all 1,340 samples across 6 brain regions and all time periods. Each point represents one sample. Samples are colored by brain region.

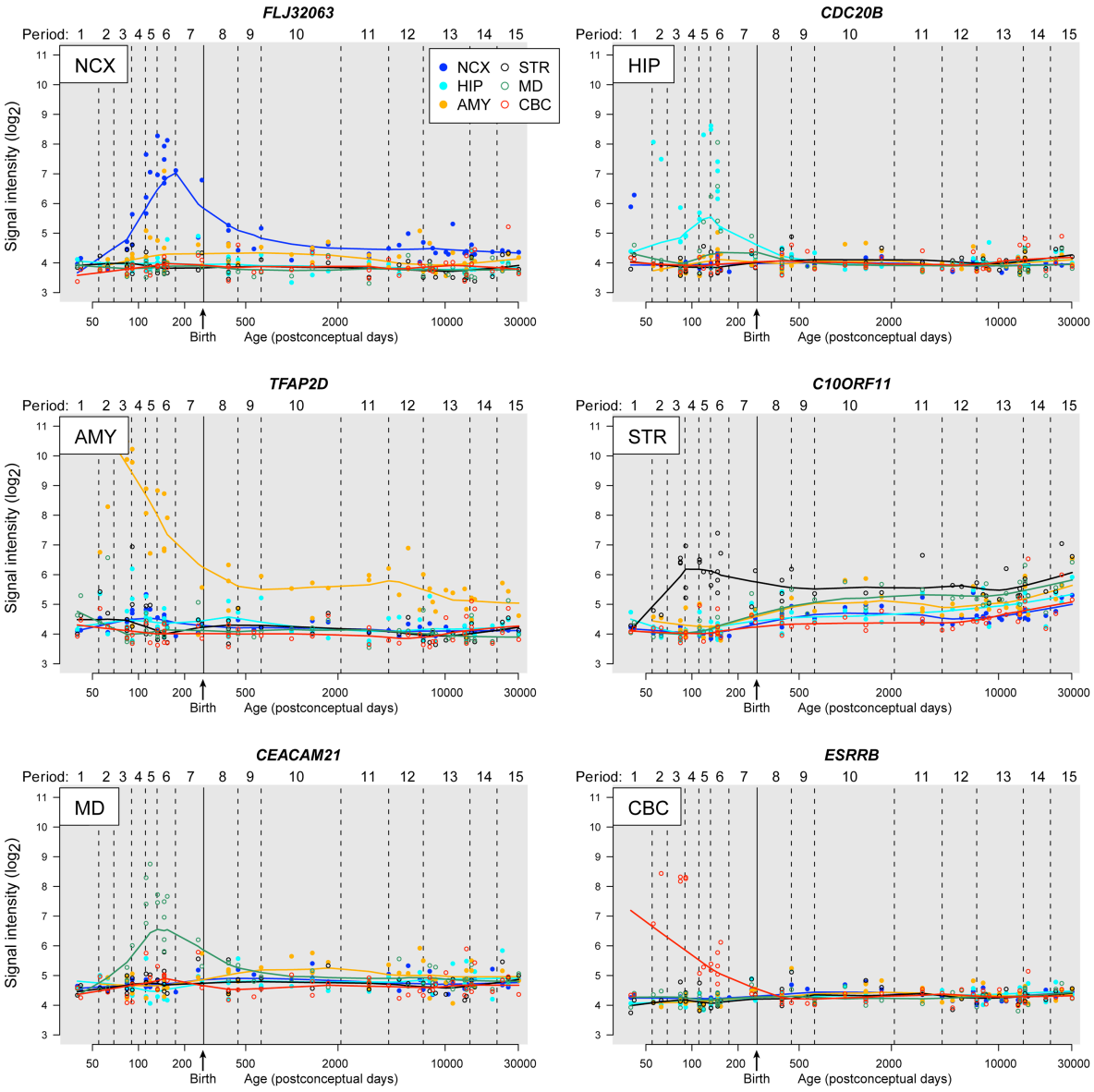


Supplementary Figure 12 | Correlations between PMI, pH, RIN, DS, and gene expression. a - c, Correlation of the number of expressed genes with PMI (a), RIN (b), and pH (c). Each data point represents one brain (a, c) or one sample (b). Weak anticorrelations and correlations were observed with PMI (Spearman correlation, *r*=-0.382) and RIN (Spearman correlation, *r*=0.278), respectively. No correlation was observed with pH (Spearman correlation, *r*=0.054). **d,** Correlation between PMI and RIN. Each data point

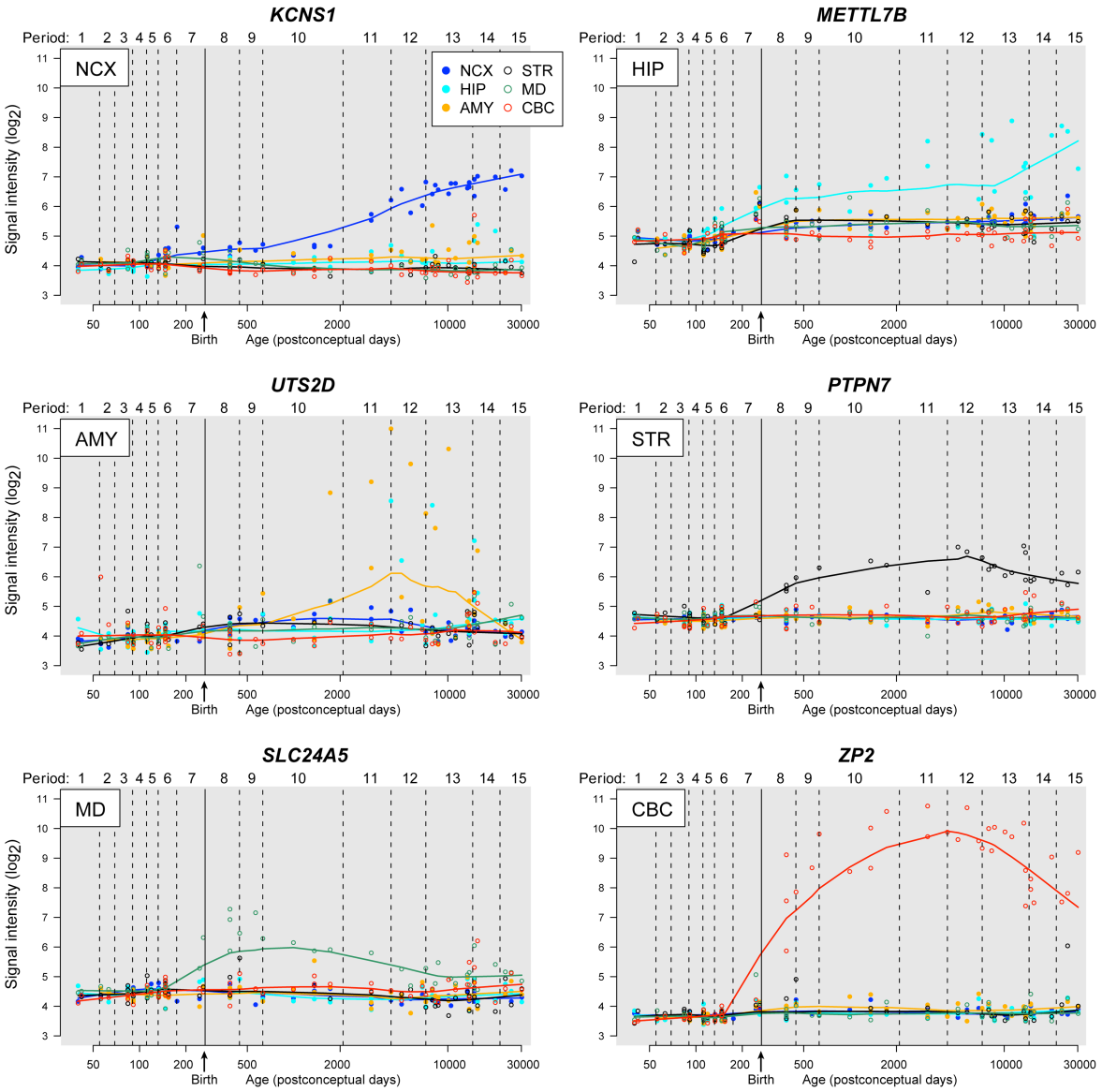
represents one brain specimen. Y-axis is the averaged RIN of individual brain specimens. X-axis is PMI of the same specimens. The Spearman correlation is -0.471, indicating that PMI and RIN are anticorrelated. **e,** The Y-axis is the number of spatial DEX genes with significantly differential expression between brain regions and NCX areas detected by ANOVA test (FDR <0.01 and 2-fold difference) for the indicated period. The X-axis is the standard deviation of DS of all samples in that period. No significant correlation was observed between variation of DS and number of spatially DEX genes (Spearman correlation, r = -0.077). We limited this analysis to periods 3-15, when regions/areas of interest are well-defined using equivalent criteria and can be consistently followed across time, thus allowing us to assess whether the impact of variation in DS within the same stage contribute to differences in the number of detected spatially DEX genes for that particular period.



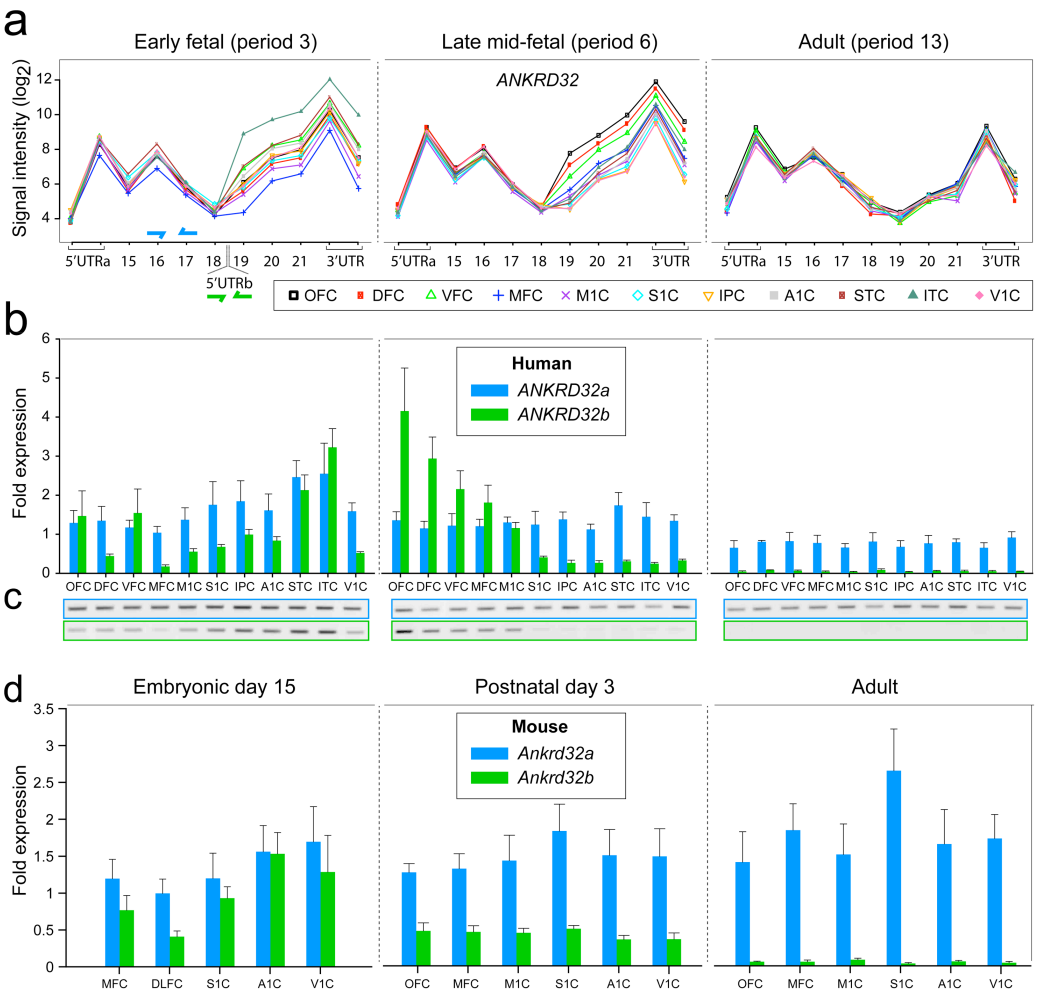
Supplementary Figure 13 | Five-fold jackknife procedure. The robustness of our DEX and DEU analysis models were tested randomly dividing samples into 5 groups and re-analyzing spatial/temporal DEX and DEU 5 times while leaving out one group of samples in each cycle. Overlap plots for DEX and DEU genes between whole dataset and each cycle are shown. The average overlap rates were as follows: spatial DEX (93.47%), temporal DEX (98.0%), spatial DEU (94.3%), and temporal DEU (99.4%). This analysis indicates that our findings did not depend on any particular sample.



Supplementary Figure 14 | Prenatal region-enriched gene expression. An example of a transiently expressed and region enriched DEX gene for each region is shown. These genes were highly expressed in during prenatal and in some cases early postnatal development in predominantly one region and then were not expressed or had low expression level in other regions. Samples and associated signal intensity values are colored by region.

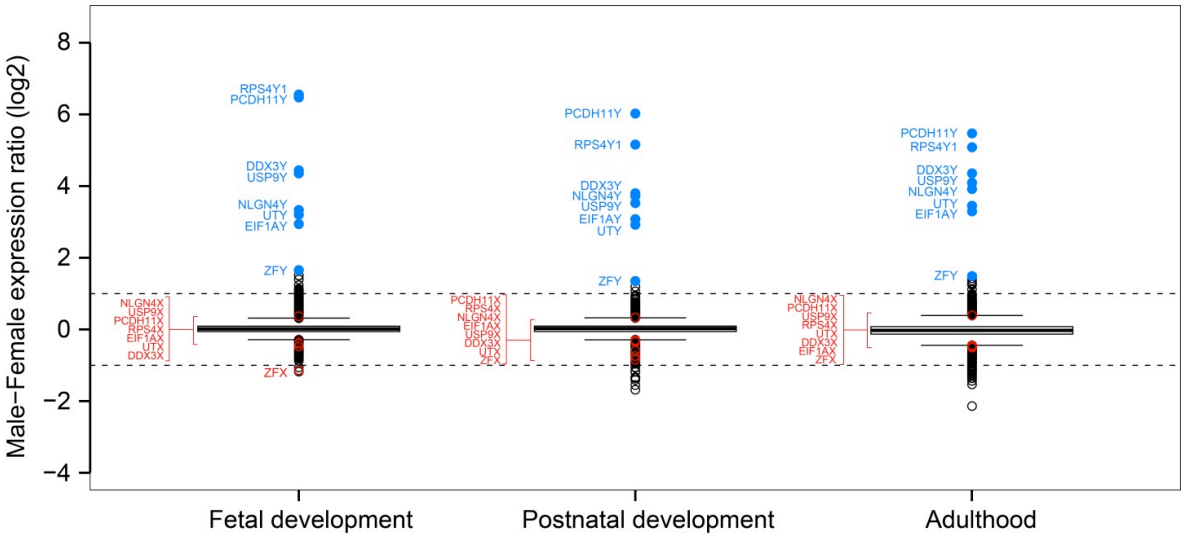


Supplementary Figure 15 | Postnatal region-enriched gene expression. An example of a transiently expressed and region enriched DEX gene for each region is shown. These genes were highly expressed in during postnatal life in predominantly one region and then were not expressed or had low expression level in other regions. Samples and associated signal intensity values are colored by region.

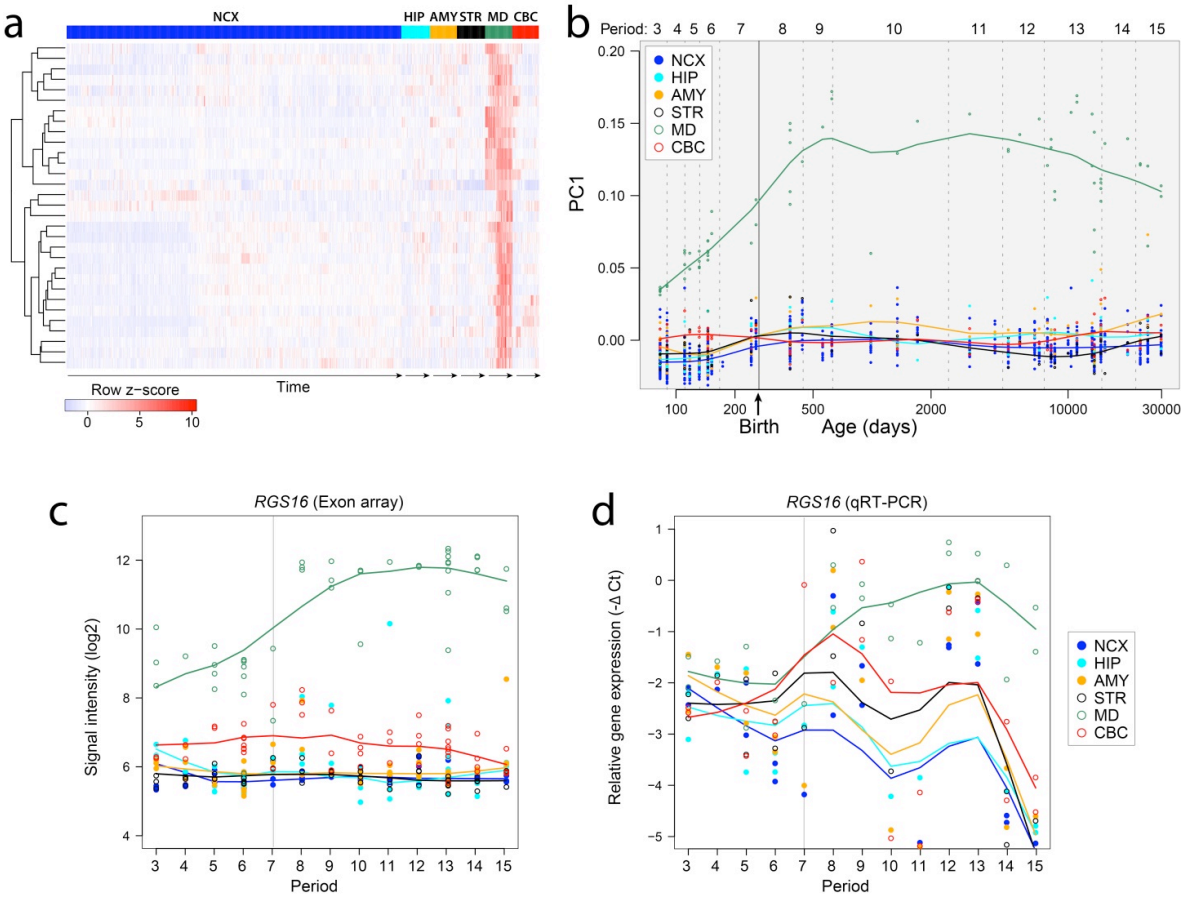


Supplementary Figure 16 | Spatiotemporal analysis of differential exon usage of *ANKRD32*. **a,** Plotted are log_2-transformed signal intensity levels (y-axis) for each probe set and the corresponding exons (x-axis) of *ANKRD32*. The exons are labeled according to RefSeq notation. The location of isoform specific primers used in b and c are depicted by blue (*ANKRD32a*) and green (*ANKRD32b*) arrows. **b,c,** qRT-PCR (**b**, mean ± SEM) and semi-quantitative RT-PCR (**c**) validation analyses of the long (blue) and short (green) *ANKRD32* isoforms (*ANKRD32a* and *ANKRD32b*, respectively) represented as relative gene expression levels. *ANKRD32a* was consistently expressed across early fetal (period 3), late midfetal (period 6), and adult NCX areas. In contrast, *ANKRD32b* exhibited significant differences in expression between areas of the early fetal and midfetal NCX (one-way ANOVA, P=1.3x10^{-5} and P=1.9x10^{-7}, respectively; followed by Tukey's pairwise test; see Supplementary Table 6). During early fetal period 3, *ANKRD32b* was most highly enriched in the ITC and, to a lesser extent, in the STC (one-way ANOVA P=1.3x10^{-5}, followed by Tukey's pairwise test). **d,** qRT-PCR in orthologous areas of the mouse NCX at equivalent developmental periods (see Supplementary Information 8) Primers were generated against the mouse genomic region orthologous to the human A*NKRD32b* 5'UTR. *Ankrd32a* and *Ankrd32b* isoforms were expressed as early as E15, which equates

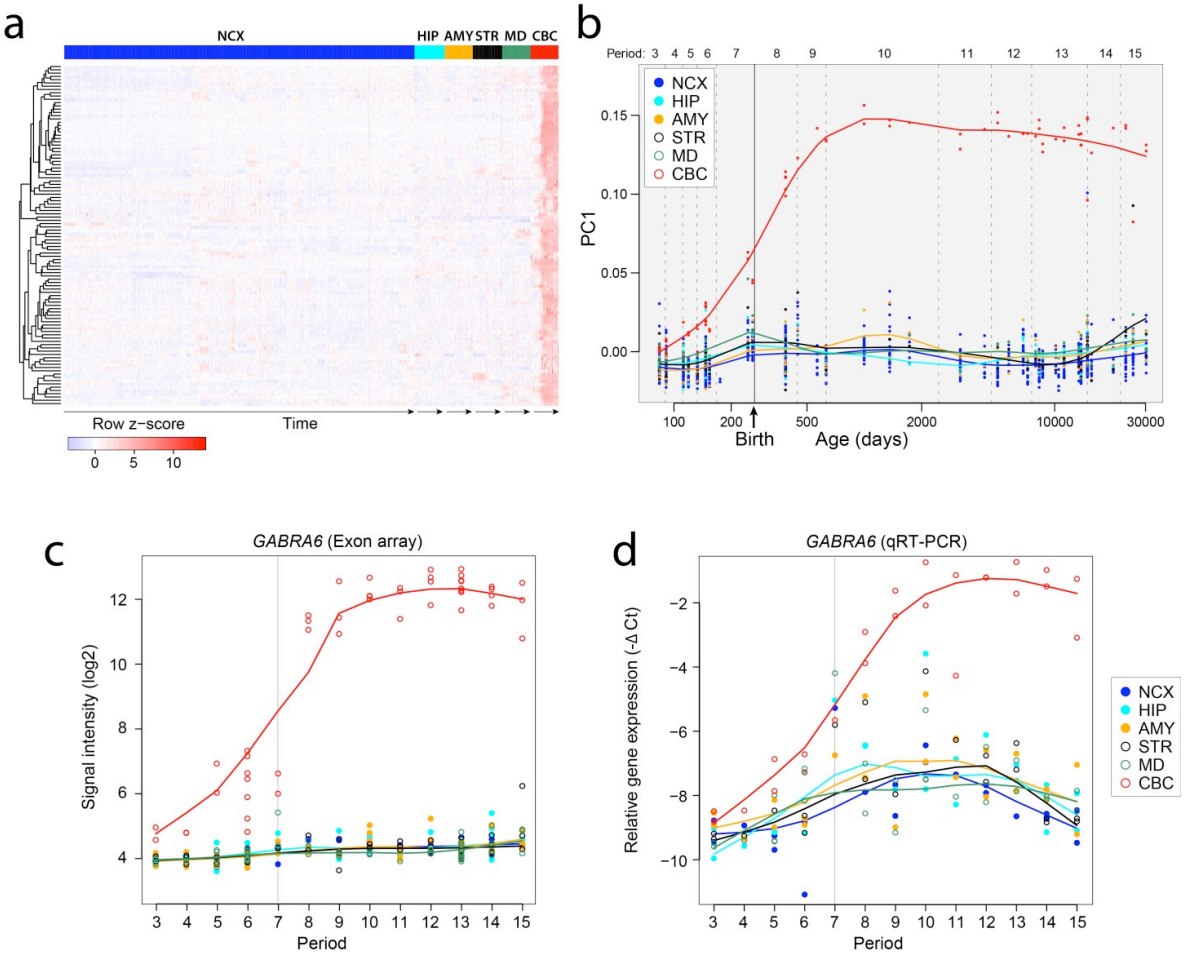
to approximately 12–13 PCW (period 3) in humans. *Ankrd32b* was expressed at lower levels compared to the *Ankrd32a* and was enriched in the prospective S1C, A1C, and V1C (see Supplementary Information 8 for tissue sampling) but not in the frontal cortex (one-way ANOVA, P=9.2x10^{-4}, followed by Tukey's pairwise test). At P3, which approximately corresponds to 20 PCW (period 6) in humans, mouse *Ankrd32b* was expressed at low levels uniformly across NCX and was not enriched in any of the NCX areas, including the multiple frontal areas (one-way ANOVA, P=0.14). We detected substantial levels of the *Ankrd32a* but not the *Ankrd32b* in the adult mouse NCX, which was similar to what was observed in the adult human NCX.



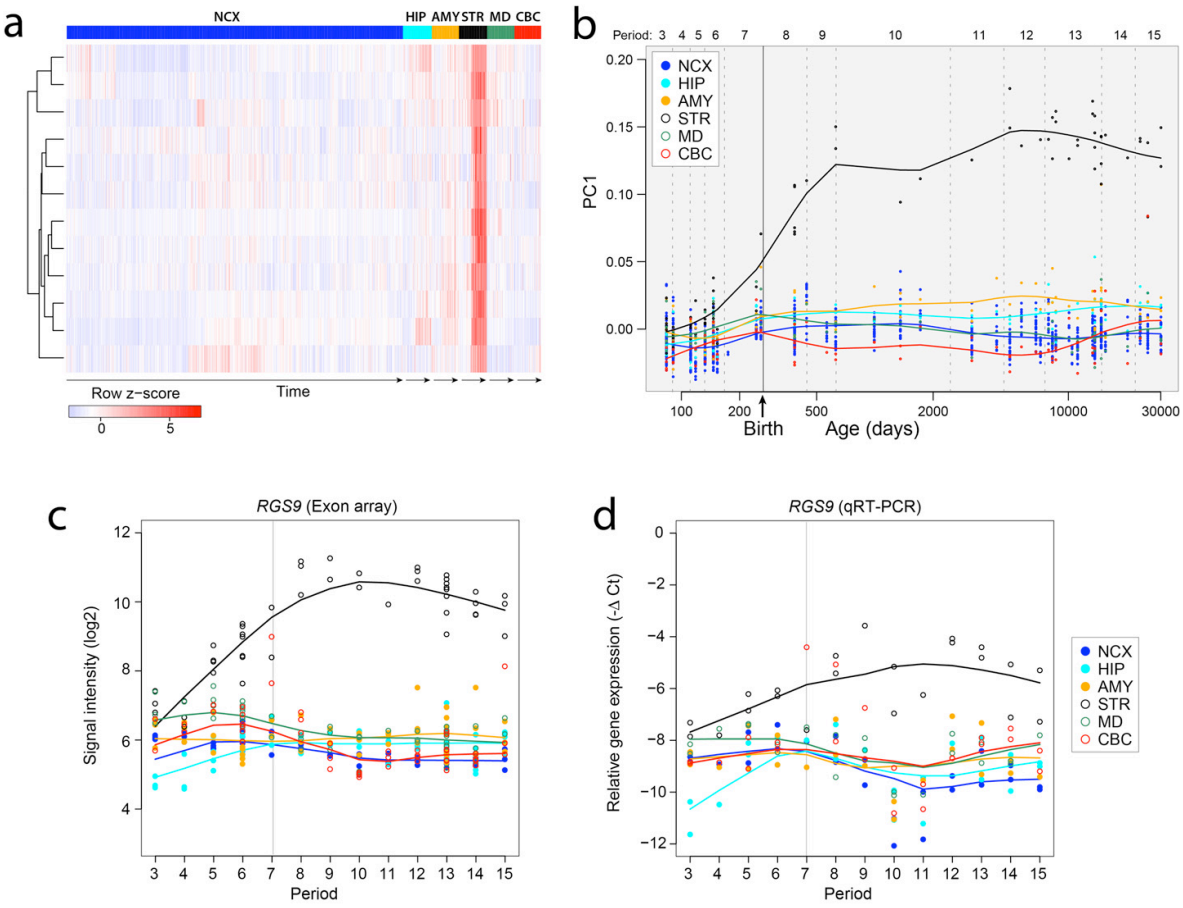
Supplementary Figure 17 | Expression of Y and X chromosome homologues in male and female brain. Ratio of male vs female median expression of protein-coding genes in all analyzed regions and NCX of male and female brain during fetal development (periods 3-7), postnatal development (periods 8-12), and adulthood (periods 13-15). One circle represents one protein-coding gene surveyed in this study. The largest differences were attributable to Y chromosome genes *PCDH11Y*, *RPS4Y1*, *USP9Y*, *DDX3Y*, *NLGN4Y*, *UTY*, *EIF1AY*, and *ZFY* (blue circles), which displayed a robust and constant expression differences across development and adulthood. Interestingly, with the exception of zinc finger X-chromosomal protein (*ZFX*), we observed that their functional homologues on the X chromosome (*PCDH11X*, *RPS4X*, *USP9X*, *DDX3X*, *NLGN4X*, *UTX*, and *EIF1AX*) were expressed at comparable levels in male and female brains across different regions and periods (Supplementary Table 7), indicating that in general these X chromosome homologues are not upregulated in a compensatory manner in developing or adult female brains.



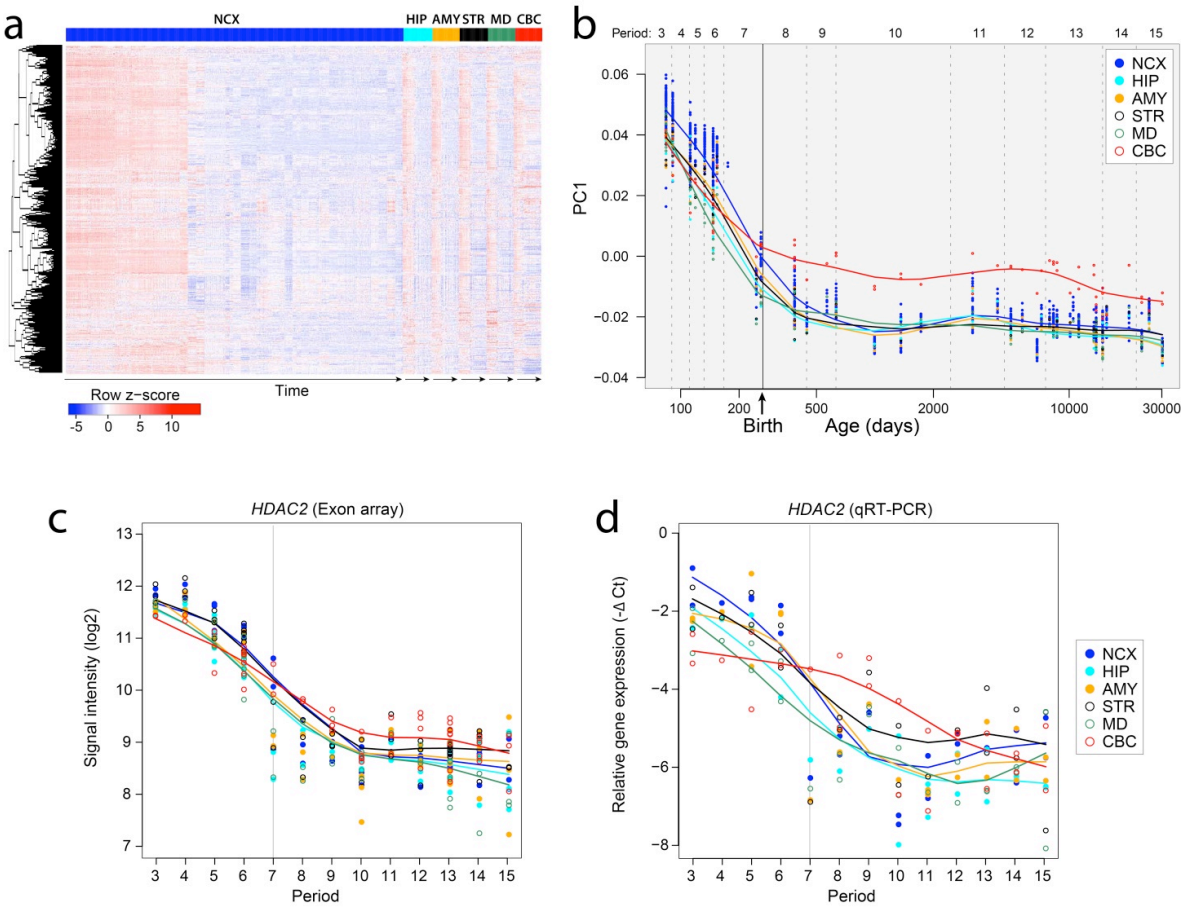
Supplementary Figure 18 | WGCNA Module 9. M9 was associated with a specific enrichment in the MD. **a,** Heat map of genes in M9 after hierarchical clustering showing the spatiotemporal pattern of the module. The expression values for each gene were ordered first by brain regions, then by age, and last by NCX areas. **b,** The spatiotemporal pattern of M9 was summarized using PCA analysis. The first principal component (PC1) was plotted against age, after being grouped and color-coded according to brain regions. The pattern was summarized by the smoothed curves of PC1 values. Dashed lines represent division between periods of the development, and the solid line separates prenatal from postnatal periods. **c, d,** Analysis of spatiotemporal expression of a representative gene, *RGS16*, with high intramodular connectivity revealed a similar pattern to the one observed for the entire module. Line plots show the $\log_2$-transformed exon array signal intensity (c) and relative expression level of quantitative RT-PCR (-ΔCt) during periods 3-15 (d).



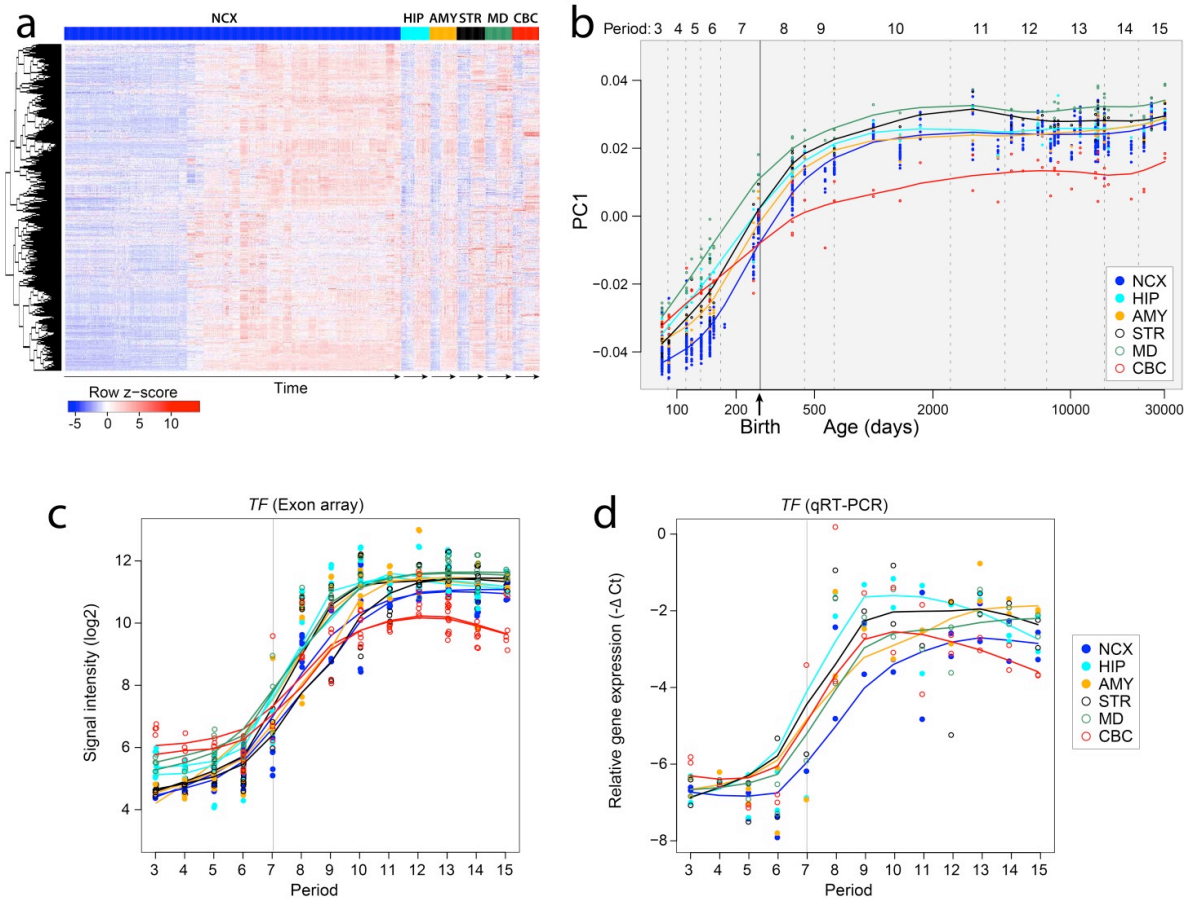
Supplementary Figure 19 | WGCNA Module 19. M19 was associated with a specific enrichment in the CBC. **a,** Heat map of genes in M19 after hierarchical clustering showing the spatiotemporal pattern of the module. The expression values for each gene were ordered first by brain regions, then by age, and last by NCX areas. **b**, The spatiotemporal pattern of M19 was summarized using PCA analysis. The first principal component (PC1) was plotted against age, after being grouped and color-coded according to brain regions. The pattern was summarized by the smoothed curves of PC1 values. Dashed lines represent division between periods of the development, and the solid line separates prenatal from postnatal periods. **c**, **d,** Analysis of spatiotemporal expression of a representative gene, *GABRA6*, with high intramodular connectivity revealed a similar pattern to the one observed for the entire module. Line plots show the $\log_2$-transformed exon array signal intensity (c) and relative expression level of quantitative RT-PCR (-ΔCt) during periods 3-15 (d).



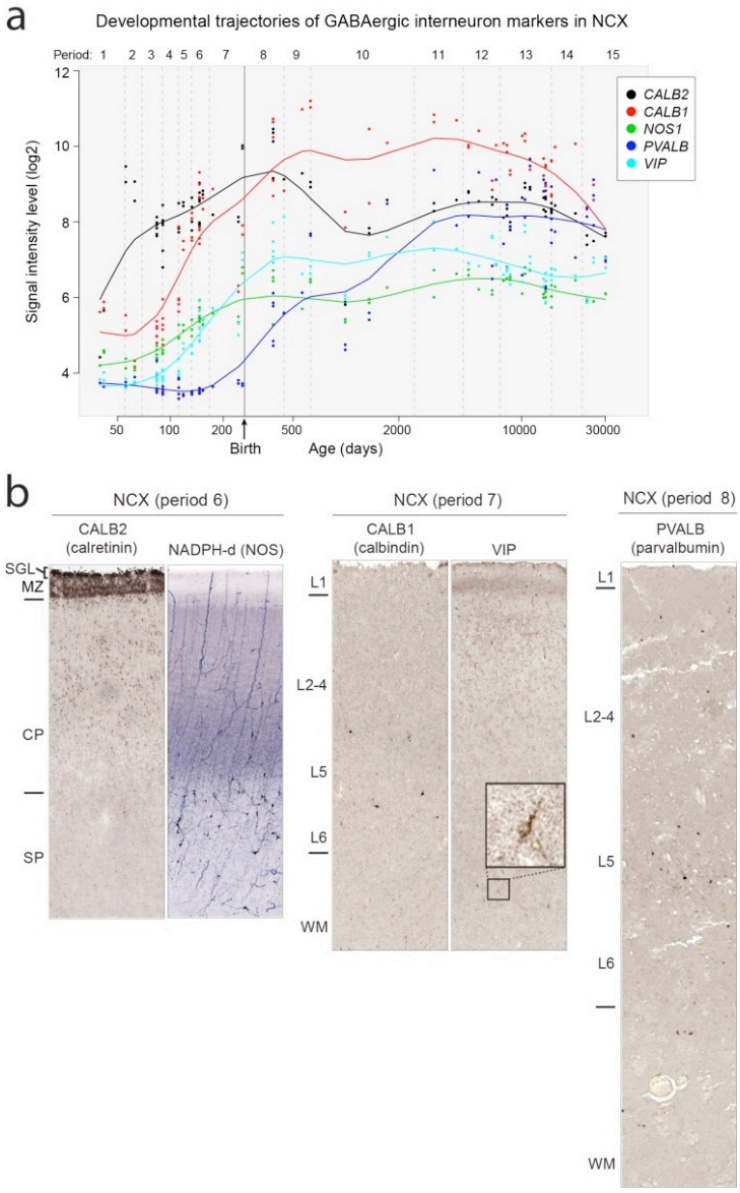
Supplementary Figure 20 | WGCNA Module 23. M23 was associated with a specific enrichment in the STR. **a,** Heat map of genes in M23 after hierarchical clustering showing the spatiotemporal pattern of the module. The expression values for each gene were ordered first by brain regions, then by age, and last by NCX areas. **b,** The spatiotemporal pattern of M23 was summarized using PCA analysis. The first principal component (PC1) was plotted against age, after being grouped and color-coded according to brain regions. The pattern was summarized by the smoothed curves of PC1 values. Dashed lines represent division between periods of the development, and the solid line separates prenatal from postnatal periods. **c**, **d,** Analysis of spatiotemporal expression of a representative gene, *RGS9*, with high intramodular connectivity revealed a similar pattern to the one observed for the entire module. Line plots show the $\log_2$-transformed exon array signal intensity (c) and relative expression level of quantitative RT-PCR (-ΔCt) during periods 3-15 (d).



Supplementary Figure 21 | WGCNA Module 20. M20 was associated with a progressive decrease in gene expression across all regions starting from the embryonic period. **a,** Heat map of genes in M20 after hierarchical clustering showing the temporally co-expressed pattern is consistent across all regions. The expression values for each gene were ordered first by brain regions, then by age, and last by NCX areas. **b,** The spatiotemporal pattern of M20 was summarized using PCA analysis. The first principal component (PC1) was plotted against age, after being grouped and color-coded according to brain regions. The pattern was summarized by the smoothed curves of PC1 values. Dashed lines represent division between periods of the development, and the solid line separates prenatal from postnatal periods. **c, d,** Analysis of spatiotemporal expression of a representative gene, *HDAC2*, with high intramodular connectivity revealed a similar pattern to the one observed for the entire module. Line plots show the $\log_2$-transformed exon array signal intensity (c) and relative expression level of quantitative RT-PCR (-ΔCt) during periods 3-15 (d).

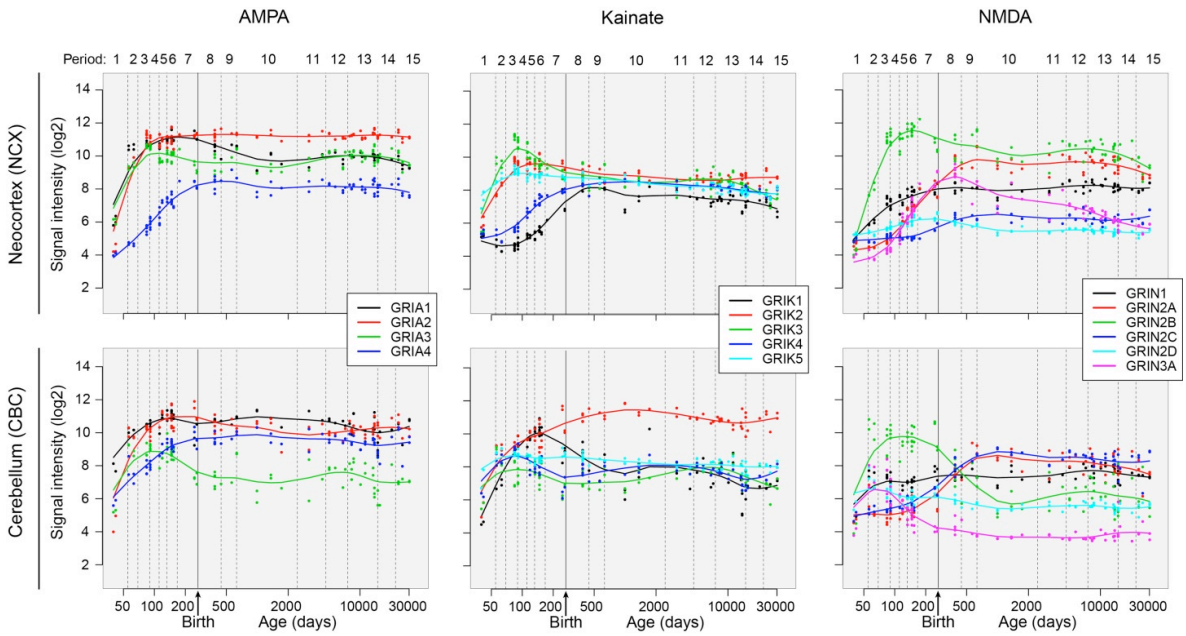


Supplementary Figure 22 | WGCNA Module 2. M2 was associated with a progressive increase in gene expression across all regions starting at the embryonic period. **a,** Heat map of genes in M2 after hierarchical clustering showing the temporally co-expressed pattern is consistent across all regions. The expression values for each gene were ordered first by brain regions, then by age, and last by NCX areas. **b,** The spatiotemporal pattern of M2 was summarized using PCA analysis. PC1 was plotted against age, after being grouped and color-coded according to brain regions. The pattern was summarized by the smoothed curves of PC1 values. Dashed lines represent division between periods of the development, and the solid line separates prenatal from postnatal periods. **c, d,** Analysis of spatiotemporal expression of a representative gene, *TF*, with high intramodular connectivity revealed a similar pattern to the one observed for the entire module. Line plots show the $\log_2$-transformed exon array signal intensity (c) and relative expression level of quantitative RT-PCR (-ΔCt) during periods 3-15 (d).



Supplementary Figure 23 | Expression trajectories of genes associated with specific cortical GABAergic interneuron subclasses. a, Transcriptome-based expression trajectories for commonly used markers of different subclasses of cortical GABAergic inhibitory neurons. **b,** Representative images of immunohistochemical detection of CALB2 (calretinin; Swant 6B3; 1:2000), NADPH-d (a histochemical marker of NOS, including NOS1), CALB1 (calbindin; Swant 300, 1:1000), VIP (vasoactive intestinal peptide; Abcam ab8795, 1:20), and PVALB (parvalbumin; Swant PV235, 1:5000) in the NCX during periods 6-8. Expression trajectories were reminiscent of the changes in the immunohistochemical detection of interneuronal markers in the fetal and early postnatal human NCX. Of the cortical GABAergic interneuron markers analyzed in this study, *CALB2* gene expression was higher than the other markers during periods 1 and 2 (black

line in a). Consistently, CALB2-immunopositive interneurons were the most abundant of the analyzed marker of GABAergic interneuron cell types in the NCX during midfetal periods (**b**). Notably, CALB2-immunopositive cells were numerous in the upper part of the cortical plate (CP). Conversely, NADPH-d/NOS1-positive interneurons were less abundant than CALB2-immunopositive interneurons (green line in a). Although some NOS1-positive cells were found in deeper layers of the CP, the majority was found in the CP-subplate (SP) border. Consistent with the gene expression data and previous independent immunohistochemical studies[73], CALB1- and VIP-immunopositive interneurons were the next to be detected in the NCX and can be easily identified during period 7. *PVALB* expression (dark blue line in a) and immunohistochemical staining (b) were the last to be detected in the NCX around birth. (SGL; subpial granular layer, MZ; marginal zone, L; layer, WM; white matter).



Supplementary Figure 24 | Expression trajectories of genes encoding subunits of glutamate receptors. Expression levels of different subunits of ionotropic glutamatergic receptors (AMPA, Kainate and NMDA) are shown for neocortex and cerebellum. Most subunits started to be expressed early during development and remained expressed throughout entire lifespan. However, spatiotemporal differences in expression patterns of different subunits were observed. Note that *GRIN3A* in CBC was the only subunit that appeared not to be expressed in postnatal or adult period.

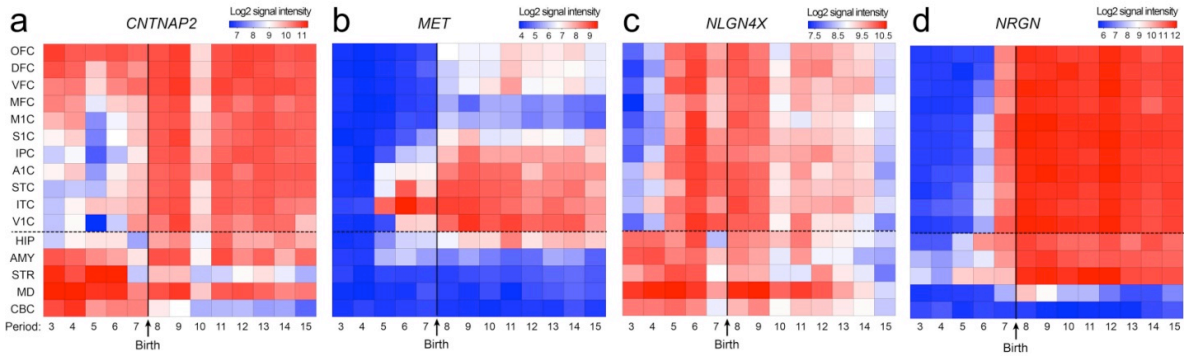


Figure 25 | Spatiotemporal expression patterns of disease-related genes. a-d, Heat map matrix representations of spatiotemporal expression of representative genes whose variants have been previously linked to (a-c) ASD (*CNTNAP2, MET, NLGN4AX*) and (d) schizophrenia (*NRGN*). The expression images display log_2-transformed signal intensity across analyzed regions/areas and time periods using a heat map color scale from low (blue) to high (red). The dashed horizontal line separates NCX areas from other brain regions. The birth time is marked by vertical solid line. **a**, *CNTNAP2*, which encodes a neurexin family protein implicated in ASD, was highly enriched in the areas of the fetal OFC and DFC, as previously reported[15,17]. The expression levels of *CNTNAP2* suddenly increased in other cortical areas during early infancy (period 8), after which it remained expressed in all NCX areas. **b**, *MET*, which encodes a hepatocyte growth factor receptor, has been linked to ASD[74] and exhibited a partially overlapping but distinct NCX areal expression pattern. *MET* was highly enriched in the early midfetal ITC and then increased in the surrounding temporal-occipital areas and the OFC, where it remained enriched throughout development. **c**, *NLGN4X*, a gene encoding a protein involved in synapse formation and function[75], was upregulated during the second half of the prenatal NCX development, which coincided with an increase in the expression of genes associated with synapse and dendrite development (Fig. 5 b, c). *NLGN4X* was also identified as a sex-specific DEU gene (Fig. 3). Mutations in *NLGN4X* have previously been associated with ASD, which are more prevalent in males. Many of the ASD-associated risk SNPs and deletions were present in exons 5, 6, and 7 *NLGN4X*[76,77], which displayed male-biased DEU. **c**, *NRGN*[40], which encodes a postsynaptic protein kinase substrate, was highly expressed in all NCX areas starting after birth, progressively increased in expression until late childhood, and remained well-expressed in adulthood.

1.10. References

1 Kostovic, I. & Judas, M. Prolonged coexistence of transient and permanent circuitry elements in the developing cerebral cortex of fetuses and preterm infants. *Dev Med Child Neurol* **48**, 388-393, (2006).

2 Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci* **10**, 724-735, (2009).

3 Rubenstein, J. L. Research Review: Development of the cerebral cortex: implications for neurodevelopmental disorders. *J Child Psychol Psychiatry*, (2010).

4 Hill, R. S. & Walsh, C. A. Molecular insights into human brain evolution. *Nature* **437**, 64-67, (2005).

5 Preuss, T., Cáceres, M., Oldham, M. & Geschwind, D. Human brain evolution: insights from microarrays. *Nat Rev Genet* **5**, 850-860, (2004).

6 Kriegstein, A., Noctor, S. & Martínez-Cerdeño, V. Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat Rev Neurosci* **7**, 883-890, (2006).

7 Lewis, D. A. & Levitt, P. Schizophrenia as a disorder of neurodevelopment. *Annu Rev Neurosci* **25**, 409-432, (2002).

8 Meyer-Lindenberg, A. & Weinberger, D. R. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nat Rev Neurosci* **7**, 818-827, (2006).

9 Insel, T. Rethinking schizophrenia. *Nature* **468**, 187-193, (2010).

10 State, M. The genetics of child psychiatric disorders: focus on autism and Tourette syndrome. *Neuron* **68**, 254-269, (2010).

11 Jamain, S. *et al.* Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet* **34**, 27-29, (2003).

12 Vawter, M. *et al.* Gender-specific gene expression in post-mortem human brain: localization to sex chromosomes. *Neuropsychopharmacology* **29**, 373-384, (2004).

13 Weickert, C. *et al.* Transcriptome analysis of male-female differences in prefrontal cortical development. *Mol Psychiatry* **14**, 558-561, (2009).

14 Reinius, B. & Jazin, E. Prenatal sex differences in the human brain. *Mol Psychiatry* **14**, 987, 988-989, (2009).

15 Abrahams, B. *et al.* Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 17849-17854, (2007).

16 Sun, T. & Walsh, C. Molecular approaches to brain asymmetry and handedness. *Nat Rev Neurosci* **7**, 655-662, (2006).

17 Johnson, M. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494-509, (2009).
18 Somel, M. *et al.* MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res* **20**, 1207-1218, (2010).
19 Ip, B. *et al.* Investigating gradients of gene expression involved in early human cortical development. *J Anat* **217**, 300-311, (2010).
20 Licatalosi, D. D. & Darnell, R. B. Splicing regulation in neurologic disease. *Neuron* **52**, 93-101, (2006).
21 Blencowe, B. J. Alternative splicing: new insights from global analyses. *Cell* **126**, 37-47, (2006).
22 Svenningsson, P. *et al.* Alterations in 5-HT1B receptor function by p11 in depression-like states. *Science* **311**, 77-80, (2006).
23 Chen, D. Y. *et al.* A critical role for IGF-II in memory consolidation and enhancement. *Nature* **469**, 491-497, (2011).
24 Lehtinen, M. K. *et al.* The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. *Neuron* **69**, 893-905, (2011).
25 Huffaker, S. J. *et al.* A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia. *Nat Med* **15**, 509-518, (2009).
26 Joutel, A. *et al.* Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* **383**, 707-710, (1996).
27 Ewart, A. K. *et al.* Hemizygosity At The elastin locus in a developmental disorder, Williams-syndrome. *Nature Genet.* **5**, 11-16, (1993).
28 Südhof, T. C. Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455**, 903-911, (2008).
29 Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* **4**, Article17, (2005).
30 Hevner, R. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29**, 353-366, (2001).
31 Molyneaux, B. J., Arlotta, P., Hirata, T., Hibi, M. & Macklis, J. D. Fezl is required for the birth and specification of corticospinal motor neurons. *Neuron* **47**, 817-831, (2005).
32 Chen, B., Schaevitz, L. & McConnell, S. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 17184-17189, (2005).
33 Chen, J., Rasin, M., Kwan, K. & Sestan, N. Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral

cortex. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 17792-17797, (2005).

34 Ariani, F. *et al.* FOXG1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet* **83**, 89-93, (2008).

35 Kwan, K. *et al.* SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 16021-16026, (2008).

36 Knoth, R. *et al.* Murine features of neurogenesis in the human hippocampus across the lifespan from 0 to 100 years. *PLoS One* **5**, e8809, (2010).

37 Han, W. *et al.* TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 3041-3046, (2011).

38 McKenna, W. L. *et al.* Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development. *J Neurosci* **31**, 549-564, (2011).

39 Perroud, N. *et al.* Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project. *Pharmacogenomics J*, doi:tpj201070 [pii] 10.1038/tpj.2010.70 (2010).

40 Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* **460**, 744-747, (2009).

41 Petanjek, Z., Judas, M., Kostović, I. & Uylings, H. Lifespan alterations of basal dendritic trees of pyramidal neurons in the human prefrontal cortex: a layer-specific pattern. *Cereb Cortex* **18**, 915-929, (2008).

42 Huttenlocher, P. R. & Dabholkar, A. S. Regional differences in synaptogenesis in human cerebral cortex. *J Comp Neurol* **387**, 167-178, (1997).

43 Stranger, B. E. *et al.* Population genomics of human gene expression. *Nat Genet* **39**, 1217-1224, (2007).

44 Heinzen, E. L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biology* **6**, e1, (2008).

45 Liu, C. *et al.* Whole-genome association mapping of gene expression in the human prefrontal cortex. *Molecular Psychiatry* **15**, 779-784, (2010).

46 Gibbs, J. R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet* **6**, e1000952, (2010).

47 Myers, A. J. *et al.* A survey of genetic human cortical gene expression. *Nat Genet* **39**, 1494-1499, (2007).

48 Webster, J. A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet* **84**, 445-458, (2009).

49 Ren, C., Ren, C. H., Li, L., Goltsov, A. A. & Thompson, T. C. Identification and characterization of RTVP1/GLIPR1-like genes, a novel p53 target gene cluster. *Genomics* **88**, 163-172, (2006).

50 Colantuoni, C. *et al.* Temporal dynamics and genetic control of transcription in the human prefrontal cortex across the lifespan. *Nature*, in press (2011).

51 O'Rahilly, R. & Müller, F. *The Embryonic Human Brain. An Atlas of Developmental Stages*. 3rd edn, (Wiley-Liss, 2006).

52 Judaš, M. in *Med Radiol Diagn Imaging* (ed D Prayer) 81-146 (Springer-Verlag, 2010).

53 Kostović, I. & Judaš, M. in *The Cognitive Neurosciences* (ed M.S. Gazzaniga) 29-47 (MIT Press, 2009).

54 Sidman, R. L. & Rakic, P. in *Histology and Histopathology of the Nervous System* (eds W. Haymaker & R.D. Adams) 3-145 (C.C. Thomas Publisher, 1982).

55 Poliakov, G. I. in *Cytoarchitectonics of the Human Cerebral Cortex* (eds S.A. Sarkisov, I.N. Filimonov, & N.S. Preobrazhenskaya) 33-91 (Medgiz, 1949).

56 Sidman, R. L. & Rakic, P. Neuronal migration, with special reference to developing human brain: a review. *Brain Res* **62**, 1-35, (1973).

57 Brodmann, K. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. (Johann Ambrosius Barth Verlag, 1909).

58 Irizarry, R. A., Ooi, S. L., Wu, Z. & Boeke, J. D. Use of mixture models in a microarray-based screening procedure for detecting differentially represented yeast mutants. *Stat Appl Genet Mol Biol* **2**, Article1, (2003).

59 Benovoy, D., Kwan, T. & Majewski, J. Effect of polymorphisms within probe-target sequences on olignonucleotide microarray experiments. *Nucleic Acids Res* **36**, 4417-4423, (2008).

60 Duan, S., Zhang, W., Bleibel, W. K., Cox, N. J. & Eileen Dolan, M. SNPinProbe_1.0: A database for filtering out probes in the Affymetrix GeneChip® Human Exon 1.0 ST array potentially affected by SNPs. *Bioinformation* **2**, 469-470, (2008).

61 Okoniewski, M. J. & Miller, C. J. Comprehensive analysis of Affymetrix exon arrays using BioConductor. *PLoS Comput Biol* **4**, 1-6, (2008).

62 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300, (1995).

63 Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* **4**, 44-57, (2008).

64 Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research* **37**, 1-13, (2009).

65 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559, (2008).

66 Hu, Z. J., Hung, J. H., Wang, Y., Chang, Y. C., Huang, C. L., Huyck, M., & DeLisi, C. VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology. *Nucleic Acids Research* **37**, W115-W121, (2009).

67 Kumar, R. A. & Christian, S. L. Genetics of autism spectrum disorders. *Curr Neurol Neurosci Rep* **9**, 188-197, (2009).

68 Allen, N. C. *et al.* Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nat Genet* **40**, 827-834, (2008).

69 Fertuzinhos, S. *et al.* Selective depletion of molecularly defined cortical interneurons in human holoprosencephaly with severe striatal hypoplasia. *Cereb Cortex* **19**, 2196-2207, (2009).

70 Clancy, B. *et al.* Web-based method for translating neurodevelopment from laboratory species to humans. *Neuroinformatics* **5**, 79-94, (2007).

71 Paxinos, G., Halliday, G., Watson, C., Koutcherov, Y. & Wang, H. *Atlas of the Developing Mouse Brain* (Academic Press – Elsevier, 2007).

72 Schambra, U. *Prenatal Mouse Brain Atlas* (Springer, 2008).

73 Meyer, G. Genetic control of neuronal migrations in human cortical development. *Adv Anat Embryol Cell Biol* **189**, 1 p preceding 1, 1-111, (2007).

74 Campbell, D. B. *et al.* A genetic variant that disrupts MET transcription is associated with autism. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16834-16839, (2006).

75 Südhof, T. C. Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455**, 903-911, (2008).

76 Lawson-Yuen, A., Saldivar, J. S., Sommer, S. & Picker, J. Familial deletion within NLGN4 associated with autism and Tourette syndrome. *Eur J Hum Genet* **16**, 614-618, (2008).

77 Qi, H. B. *et al.* Positive association of neuroligin-4 gene with nonspecific mental retardation in the Qinba Mountains Region of China. *Psychiatr Genet* **19**, 1-5, (2009).

2. Temporal Dynamics and Bilaterality of Human Neocortical Areal Transcription

Mihovil Pletikos[1,2]†, André M. M. Sousa[1,3]†, Goran Sedmak[1,2]†, Kyle A. Meyer[1], Ying Zhu[1], Feng Cheng[1,4], Nenad Šestan[1*]

[1] Department of Neurobiology and Kavli Institute for Neuroscience, Yale School of Medicine, New Haven, Connecticut 06510, USA.

[2] Graduate Program in Neuroscience, Croatian Institute for Brain Research, University of Zagreb School of Medicine, 10000 Zagreb, Croatia.

[3] Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical Sciences Institute, University of Porto, 4099-003 Porto, Portugal

[4] Department of Pharmaceutical Sciences, College of Pharmacy, University of South Florida, Tampa, FL, 33612, USA.

†These authors contributed equally to this work.

*To whom correspondence should be addressed. E-mail: nenad.sestan@yale.edu

2.1. Abstract

Transcriptional processes involved in the areal patterning and lateralization of the human neocortex are poorly understood. Here, we analyzed the temporal dynamics and laterality of protein-coding gene expression across topographically defined areas of the prenatal and postnatal human neocortex. We found that inter-areal transcriptional differences exhibit a temporal hourglass pattern, characterized by robust expression differences prenatally and, to a lesser extent, from adolescence onward, though not in infancy and childhood. Further analyses of temporally regulated inter-areal divergence revealed distinct gradients and/or compartments of expression, including some that were different between human and rhesus macaque. We also found that global gene expression trajectories of each area become increasingly synchronized in postnatal development. Analyses of gene expression between left and right areas revealed global symmetry throughout the fetal development and postnatal lifespan. Thus, human neocortical areal transcription is temporally specified, increasingly synchronized over the course of postnatal development, and globally symmetric.

2.2. Introduction

The human cerebral neocortex (NCX) is parcellated into functionally distinct sensory, motor, and association areas that provide biological substrates underlying perception, behavior, and cognition (*1-7*). The functional specializations of these modalities within a neocortical area depend on the pattern of its synaptic connections. While the basic features of this areal map and connections are shared among mammals, important variations exist in their organization that provides a substrate for human-specific cognition and behavior (*8-13*). For example, when compared with the rhesus macaque, a closely related and the most studied non-human primate (NHP), the human NCX has approximately an 8-fold larger surface area and more functionally distinct areas (*6*). Yet this areal expansion was not uniform, as it seemed to have predominantly affected frontal and temporo-parietal association areas (*6, 9, 11-13*).

Another key feature of the human NCX is that it covers the surface of the left and right hemispheres, each comprising a topographically matched, though slightly structurally and functionally asymmetric areal map (*14-25*). This asymmetric organization plays a crucial role in functional lateralization of many cognitive and motor functions, such as language and handedness, between the hemispheres. Multiple lines of evidence indicate that these asymmetries are reflected at the molecular (*21*), cellular (*22, 23*) and circuitry level (*18, 25*). First signs of structural asymmetry appear during the late mid-fetal period (*20, 24, 25*) and become more prominent during early postnatal development when functional asymmetries become noticeable (*25*).

Several lines of evidence indicate that distinct human neocortical areas, as well as hemispheres as a whole, mature at different rates (*26-30*). For example, axons in primary sensory-motor areas start to myelinate before those in the association areas (*26*). Other processes such as synaptogenesis also exhibit prominent inter-areal differences in their maturational trajectories (*27*). Finally, there is also evidence indicating that the right hemisphere matures faster than the left during late fetal and early postnatal development (*31-33*).

There is increasing evidence to suggest that processes regulating both areal patterning and asymmetry, as well as the maturational trajectories of these processes, are affected in major psychiatric and neurological disorders (25, *30, 33-36*). Moreover, the progression of certain neuropathologies follows a stereotypic areal pattern (*37*), indicating that the mechanisms involved in the areal patterning and asymmetry may play a role in disease manifestation. However, little is known about these developmental processes in normal and diseased brains, or how they differ from other mammals, especially closely related NHPs.

The majority of what we know about the areal patterning of NCX has been derived from experimental studies in a handful of mammals, mostly through work in the mouse. These studies have revealed that the positional identity of areas is defined by the combination of patterning genes with distinct spatio-temporal expression patterns and extrinsic influence of ingrowing corticopetal axons, in particular those originating from the thalamus (*2-5*). However, the relevance of these studies to human NCX development is limited for several reasons. In addition to its increased size and complexity compared to the rodent NCX (*10, 12*), the development of the human NCX is spread over a much longer time period (*9, 12, 38*), so that even those rodent studies that sample multiple time points are difficult to relate to specific developmental processes in the human. In contrast to areal patterning, the development of NCX asymmetry has been studied almost exclusively in humans, but the number of such studies has been very limited. A number of genes were found to be expressed asymmetrically during the early fetal (*21*), but not the late mid-fetal or adult NCX (*39, 40*), suggesting that transcriptional asymmetry may be temporally regulated and precede structural and functional lateralization. In the present study, we sought to substantially extend on this previous research, both qualitatively and quantitatively, by analyzing the transcriptome of a spatially and temporally rich human neocortical collection, and to compare some of the transcriptome findings with the rhesus macaque.

2.3. Results

Inter-areal transcriptional divergence exhibits a temporal hourglass pattern.

To gain insight into the transcriptional processes involved in the areal patterning of the human NCX, we performed a secondary analysis of exon-level mRNA transcriptome data of 11 topographically defined areas (*38*): the orbital (OFC), dorsolateral (DFC), ventrolateral (VFC), medial (MFC) and primary motor (M1C) areas of the frontal lobe; the primary somatosensory (S1C) and posterior inferior (IPC) areas of the parietal lobe; the primary auditory (A1C), posterior superior (STC) and anterior inferior (ITC) areas of the temporal lobe; and the primary visual (V1) area of the occipital lobe (Fig. 1A and Supplementary Material and Information). This dataset was generated using 886 tissue samples isolated from left and right hemispheres of 53 clinically unremarkable post-mortem human brain specimens spanning from early fetal development through ageing (from 10 weeks of post-conception (PCW) to 82 years of age) or periods 3-15 as previously designated (*38*), an interval during which the areas can be consistently followed across time.

Assayed tissue samples showed robust and unique transcriptional signatures. Hierarchical clustering based on all transcriptome data revealed grouping of areas by their

topographical proximity and functional segregation (Fig. 1A). However, principal component analysis (PCA) showed that transcriptional differences across periods account for the majority of the variance among samples (fig. S1). Analysis of gene expression using the core probe sets (Supplementary Note 3.1) revealed that 13,956 of the 17,557 assayed genes (74.49%) were expressed in the NCX in at least one area of at least one period. Interestingly, the average number of genes expressed in all 11 NCX areas per brain remained relatively constant across time (10,833 +/- 462 s.d.). However, more genes were found to be expressed at any time or area during 28 weeks of fetal development (periods 3-7) than during the first 20 years of postnatal development (periods 8-12) or several decades of adulthood (periods 13-15) (Fig. 1B). Negligible correlations were observed between expressed genes and several possible confounding factors: the postmortem interval (PMI), RNA integrity number (RIN), and pH (figs. S2-S4). These findings indicate that gene expression is more dynamic during fetal development, likely reflecting transcriptional differences between fetal areas and periods.

To estimate the degree of transcriptional differences between areas and how these may change across time, we use analysis of variance (ANOVA) to identify genes that exhibit differential expression (DEX) among areas per period (henceforth referred to as the inter-areal divergence; Supplementary Note 3.3). We found that inter-areal transcriptional divergence exhibits a temporal hourglass pattern, characterized by robust and dynamic differences prenatally and, to a lesser extent from adolescence onwards, though not in infancy and childhood (Fig. 1C and fig. S5). To estimate the contribution of individual areas to the overall inter-areal divergence, Tukey's pairwise comparison was performed after ANOVA to determine the total number of significant ($P < 0.01$) DEX gene comparisons of each area with all the other areas for a given period. The relative contribution of areas to the overall hourglass shape varied across periods. During fetal periods, MFC, ITC, and the primary areas (V1C, A1C, S1C and M1C) exhibited the most prominent dissimilarity, whereas only MFC and V1C showed robust dissimilarities during adolescence and adulthood (Fig. 1C). Together, these findings show that the pattern of inter-areal transcriptional divergence is temporally specified and exhibits an hourglass pattern, with infancy and childhood representing a long phase of minimal divergence. Our results also show that the spatial pattern of inter-areal divergence is mainly driven by a subset of areas.

Temporal transcriptional hourglass pattern reflects areal and functional differences. We hypothesized that increased inter-areal transcriptional divergence during fetal development and from adolescence onwards reflect the differences in the underlying molecular and cellular processes between these two phases. Consistent with this

hypothesis, only 848 of 3,125 (27%) inter-areal DEX genes were DEX both in fetal development and from adolescence onwards. To gain insights into the differential organization of the NCX transcriptomes during the two phases of increased inter-areal differences, we performed weighted gene co-expression network analysis (Supplementary Note 3.7) to identify modules of co-expressed genes with often-shared functional relevance. Within fetal development, we identified 122 modules (M1-M122: table S4), and from adolescence on we found 207 modules (M123-M329; table S5). Functional annotation of the modules revealed significant differences between the organization of fetal and adolescent/adult differential areal transcriptomes. Furthermore, the gene ontology (GO) enrichment analysis (Supplementary Note 3.8) revealed significant differences between fetal and adolescent/adult DEX and intra-modular genes in the enrichment for GO categories (tables S3, S4, and S5). Exclusively fetal GO categories included: phosphoprotein, neuron differentiation, cell cycle, neuron development, cell morphogenesis, mitosis, cell morphogenesis involved in neuron differentiation, cell adhesion, and cellular component morphogenesis (Bonferroni-adjusted P ≤6.61E-11 per category). In contrast, exclusively adolescent and adult GO categories included: synaptic vesicle, transport, domain:C2 1 and 2, intrinsic to plasma membrane, intrinsic to membrane, clathrin-coated vesicle, neurological system process, neurotransmitter binding, coated vesicle, integral to membrane, and monovalent inorganic cation transport (Bonferroni-adjusted P ≤ 2.39E-06 per category).

Many fetal co-expression modules had pronounced areal gradients and/or well-defined compartments with distinct boundaries (Figs. 2A-H and fig. S13). In addition to expected gradients with prefrontal/frontal-enriched graded expression along the anterio-posterior axis (e.g., M54, M62, M80, M91), we found patterns including those with temporal (e.g., M2, M13, M5), occipital (e.g., M105), perisylvian (e.g., M118), and ventro-medial areas (MFC, OFC, V1C; e.g., M100). The majority of the modules exhibited temporally specified areal patterns that changed dramatically across fetal periods and lost their prominent areal differences postnatally (fig. S13; see also tables S4 and S5 with description of spatio-temporal patterns). Areal expression patterns of 6 selected intra-modular hub genes were validated (median correlation with array data, r = 0.88) by quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR) (Fig. 2I-L and fig. S14). *CLMP* (*ASAM*), a gene encoding an adhesion molecule (Fig. 2I and fig. S14A-B), and *WNT7B*, a gene involved in early brain development and dendritic arborization (*41*) (Fig. 2K and fig. S14E-F), displayed an expression gradient along the rostro-caudal axis in anterior-posterior and posterior-anterior directions, respectively. We also validated genes with enriched expression in the temporal areas, such as *NR2F2* (*COUP-TF2*) (Fig. 2L and fig. S14G-H), a gene enriched in the caudal ganglionic eminence and involved in

migration of interneurons (*42*), or restricted to the temporal areas, including the previously uncharacterized gene *C13ORF38* (Fig. 2J and fig. S14C-D).

Adolescence and adulthood co-expression modules exhibited a more temporally steady and less complex spatial patterns (Fig. 3A-E and fig. S15). Consistent with our previous finding of dramatic differences in the number of DEX gene among areas in these periods (Fig. 1C), the most common areal co-expression patterns reflected enrichments in MFC (e.g., M215), V1C (e.g., M214), MFC and ITC (e.g., M239), and differences between the primary sensory-motor areas (e.g., M269), and association areas (e.g., M183). Expression patterns of several intra-modular hub genes were validated with qRT-PCR (median r = 0.82), such as *GABRQ* in MFC, *TRPC3* in V1C, *VAV3* in primary sensory areas, and *BAIAP3* in all association areas (Fig. 3 and fig. S16). Together, these findings indicate that the transcriptional hourglass shape reflects temporal differences in processes related to the construction and functional specializations of areas.

Inter-areal differences exhibit inter-primate divergence.

To investigate whether some of the observed temporally regulated inter-areal expression patterns exhibit inter-primate divergence, we performed qRT-PCR on the previously selected intra-modular hub genes (Figs. 2I-L and 3F-I, and figs. S14 and S16) in homologous regions of human and rhesus macaque neocortices at equivalent ages (Supplementary Note 5). Of the hub genes that had prominent inter-areal differential expression prenatally but not postnatally (*CLMP*, *C13ORF38*, *NR2F2*, *KCNK12*, *WBSCR17* and *WNT7B*), only *NR2F2* and *WNT7B* had spatio-temporal areal expression profiles that were well correlated (r > 0.8) between human and macaque, indicating that their expression patterns are well conserved. In contrast, the inter-areal expression pattern of *WBSCR17* was only conserved during late-fetal periods *(*r > 0.8), while *KCNK12* had moderate correlation only during early fetal periods (r = 0.63), indicating that the some of the spatio-temporal expression patterns are shaped by regulatory programs that differ between humans and NHPs. Comparative analysis of adolescence and adulthood intra-modular hub genes (*BAIAP3*, *GABRQ*, *TRPC3* and *VAV3*) revealed that the inter-areal expression patterns of *BAIAP3* and *VAV3* were conserved in both adolescent and adult human and macaque NCX (r > 0.8) (Fig. 3F-I and fig. S16). In contrast, the areal expression patterns of *GABRQ* and *TRPC3* were only conserved in adulthood (r > 0.8) but not in adolescence (r ≤ 0.6). While limited in their scope, these findings show that certain intra-modular hub genes exhibit temporally regulated and species-differential inter-areal expression patterns, and may reflect broader differences in developmental gene expression programs between the two primates.

Areal transcriptome trajectories become more synchronized during postnatal development.

Previous work has shown that NCX areas mature at different rates with primary sensory-motor areas maturing first, followed by association areas (*26-30, 36*). To investigate the transcriptional dynamics associated with areal maturation, we generated intra-areal global expression trajectories by correlating the transcriptome profile of each sample to the averaged transcriptome profile of the corresponding area in mid-adulthood (period 14; Supplementary Note 3.5.1). We also correlated the transcriptome profiles to the corresponding area in early fetal development (period 3), to control for putative bias of selecting only a specific period (Fig. 4A). These trajectories allowed us to detect large-scale changes in the transcriptome over time in individual areas. Surprisingly, all areal maturational trajectories had similar shape, with steep increase during mid- and late fetal development and the major inflection point on the curve during late infancy (Fig. 4A). However, the average deviation of the areal trajectory from the average overall maturational trajectory (maturational difference index; see Supplementary Note 3.5.2) varied more among areas during fetal than postnatal periods (Fig. 4B). MFC and ITC appear to mature faster than other areas prenatally, while DFC and V1C mature slower. Thus, inter-areal differences in maturational rates did not follow global antero-posterior or medio-lateral neurogenetic gradients observed in rodents (*43*). Furthermore, these findings indicate that at the global level areal transcriptomes become more synchronized during postnatal development. Interestingly, we also observed a decline in some areal trajectories, most prominently V1C, in period 15. Moreover, V1C changes in aging deviate in the same direction in both fetal and adult trajectories suggesting that the changes are not a reversal toward fetal expression as previously reported (*44*), but instead are likely unique to the aging process. These transcriptome changes could not be explained by the difference in confounding factors between V1C and non-V1C areas (RIN, $P = 0.41$; PMI, $P = 1$; pH, $P = 1$), as well as between adult periods (RIN, $P = 0.96$; PMI, $P = 0.84$; pH, $P = 0.13$). Consistent with previous findings in aged prefrontal cortex (*44*), this suggests that aging is associated with global transcriptomic changes that vary considerably among areas.

Next, we followed the maturational trajectories of specific neurobiological processes and categories, using the first principal component to summarize the expression of genes associated with various developmental processes, neural cell types, and neurotransmitter systems (Supplementary Note 3.6). Notable trajectories and differences in their onset times, rates of increase and decrease, and shapes were observed within and between areas. Many of them showed more prominent inter-areal variations prenatally than postnatally (Supplementary Note 3.6 and figs. S6-12). These

included cell proliferation (which displayed antero-posterior and medio-later gradients), dendrite development, synapse development (which showed accelerated rate in M1C and S1C), glutamatergic neurons, astrocyte-specific genes, *CALB1* (a marker of specific interneuron type), muscarinic receptor M3, and kainite inotropic receptor genes. In contrast, it was rarer to have trajectories exhibiting more robust inter-areal variations at postnatal periods, such as the expression of *PVALB* (a marker of specific interneuron type). Some trajectories, such as myelination, were variable both prenatally and postnatally, while some, such as, oligodendrocyte- and microglia-specific genes (fig. S10C-F) remained very similar between areas across both prenatal and postnatal periods. These findings indicate that specific transcriptional trajectories vary across the areas, with prenatal development being generally more variable among areas.

Areal transcriptomes are globally symmetric across time.

Since the initial discoveries of the functional lateralization of speech and language in the perisylvian areas of NCX (*14, 45*), an abundance of data has demonstrated functional and structural asymmetries between hemispheres (*16-20, 23, 24, 33*). However, transcriptome studies of left-right differences in areal gene expression have remained rare and limited to a small temporally restricted dataset (*21, 39, 40*). Furthermore, the previous findings are also potentially inconsistent, as evidence for gene expression asymmetry has been found in early fetal (*21*), but not the mid-fetal or adult NCX (*39, 40*). Our analysis of the left-right areal transcriptomes across periods 3-15 failed to identify statistically significant inter-hemispheric areal differences, using a paired *t*-test with an FDR threshold of 0.01. To increase the dynamic range and sample size for left-right comparison, we also repeated the analysis using a sliding window algorithm that covers four consecutive periods. On average, 99.4% of expressed genes had smaller than 1.5 fold differences in left-right expression and this number varied slightly across time (Fig. 5A and B). However, none of these or the remaining 0.6% of genes with more than 1.5 fold differences in left-right expression passed the FDR threshold.

To test the concept of the arrival of corresponding developmental stages sooner in the right than in the left hemisphere (*24, 32*), we compared the individual areal maturational trajectories between hemispheres. This analysis revealed no trend in faster maturation of one hemisphere over the other (Fig. 5C and D). Our findings indicate that the well-described examples of functional and structural lateralization are not reflected in transcriptome asymmetry at the population level across almost the full course of human neocortical development and adulthood.

2.4. Discussion

Taken together, our analyses of temporal dynamics and bilaterality of areal transcriptomes revealed several novel aspects of human neocortical development. Although the total number of genes expressed in NCX remained relatively constant across time, we found the degree of transcriptional differences between areas, as measured by the number of DEX genes, is temporally regulated exhibiting an hourglass pattern. The greatest inter-areal differences were present prenatally, followed by periods from adolescence onwards, with infancy and childhood representing a phase of minimal inter-areal transcriptional divergence.

We also provided evidence that these three global phases of inter-areal transcriptional divergence reflect underlying developmental and functional differences. Fetal DEX genes and co-expression modules had highly dynamic spatio-temporal expression patterns and tended to be related to early neurodevelopmental processes associated with cell generation and neural circuit construction. Consistent with this finding, previous neuroanatomical studies have shown that area-specific features of axonal projections and neuronal differentiation are already evident in the mid-fetal NCX (*39, 46, 47*). Interestingly, the relative contribution of areas to the overall inter-areal transcriptional divergence varied across fetal development and may reflect the underlying processes involved in areal patterning. Additionally, our results indicate that fetal MFC and ITC, which have high number of DEX genes and whose global transcriptional trajectories point to faster maturation, may have differentiated earlier than other analyzed areas. This developmental pattern resembles the evolutionary pattern of NCX differentiation arising from dual archicortical and paleocortical moieties (*48*). Also, increased transcriptional divergence of prospective primary motor-sensory areas (V1C, A1C, S1C, and M1C) pointed to their early transcriptional specification.

The fetal phase is followed by a sharp drop in inter-areal transcriptional divergence during late fetal development and the appearance of a long phase covering infancy and childhood in which areas become more similar. In addition, the global gene expression maturational trajectories become more synchronized over the course of early postnatal development. We hypothesize that this areal transition from high divergence to similarity as well as increased synchronization are related to changes in neurodevelopmental processes that occur at the same time. Over one hundred years ago, Brodmann discovered (*1*) that the anatomy of the entire neocortical plate is transiently transformed by the appearance of an ontogenetic six-layered Grundtypus (i.e., the fetal equivalent of future layers 2–6) during late fetal and neonatal development (periods 7 and 8 in this study). More recently, it has been shown that during those periods layer 5 pyramidal

neurons of the whole human NCX transiently express nitric oxide synthase 1 protein (*47*). Thus, these global structural and molecular changes in the late fetal and neonatal human NCX may function as mechanisms by which areas become more similar in their expression patterns.

The drop in inter-areal transcriptional divergence also occurs while the human neocortical map is being reorganized, largely through the influence of extrinsic forces like activity-dependent mechanisms, experience, and environmental and social input (*3, 46, 49*). Also, infancy and childhood are critical and sensitive periods for the acquisition and refinement of cognitive and motor functions characterized by higher plasticity (*36, 46, 49*). Thus, we hypothesize that following the initial establishment of area-specific subcortical and cortico-cortical projections in fetal development, the robust inter-areal transcriptional differences are no longer needed and are surpassed by more general neuronal and glial differentiation transcriptional programs involved in maturation of neural circuits. Some of these more general programs, such as the transcription of myelination genes, and some interneuron and neurotransmitter systems-related genes, also exhibit inter-areal difference in the neonatal and postnatal NCX, but they appear to be less robust or less dynamically regulated than many of the fetal programs. Interestingly, specific transcriptional trajectories seem to be more synchronized between areas than expected from analyses of maturational trajectories of several phenotypic traits such as synaptogenesis (*27*), cortical thickness (*28, 29, 36*) and myelination (*26*). This finding also suggests these phenotypic inter-areal differences arise in part by post-transcriptional and activity-dependent mechanisms. Furthermore, early postnatal inter-areal transcriptional differences were more pronounced for some interneuron markers and neurotransmitter systems, suggesting that they either play a role in postnatal functional areal differentiation and/or are substantially affected by this process.

Inter-areal transcriptional divergence increases again during adolescence, a period of great developmental changes in the human brain (*49*), and slowly increases as the brain matures and ages. Unlike genes from the fetal phase, DEX genes and co-expressed genes during adolescence and adulthood tend to be related to categories such as synaptic transmission and cell-cell signaling. Thus, adolescence and adulthood, especially the aging process, are accompanied by specific changes in areal transcriptomes.

Interestingly, the temporal pattern of inter-areal transcriptional divergence resembled the hourglass model of transcriptome divergence between various species of the fruitfly (*50*). We show here that some of the genes that have temporally regulated inter-areal differential expression also display divergence, both spatially and temporally, between human and macaque monkey. These differences in gene expression patterns between species may have played a role in human evolution.

Although surprising, given well-described structural and functional asymmetries, our findings show that areal transcriptomes are largely symmetric at the population level and mature at similar rates in the left and right hemispheres across the full course of human brain development and adulthood. In accordance to this finding, a recent large-scale structural MRI study has also shown population-level cortical symmetry (*7*). With the present study design, we cannot exclude that transcriptional asymmetry is present in areas not profiled in this study, may be diluted at the population level by inter-individual variations, may be limited to the small subset of genes not present on the array, or may be driven by more subtle changes in specific cellular components that we have not been able to detect. However, it may also be that neocortical asymmetry is driven by stochastic, hormonal, epigenetic, functional connectivity-based, and/or activity-dependent mechanisms that are not clearly reflected in gene expression.

2.5. Methods Summary

The Supplementary Information provides a full description of materials and methods, including tissue acquisition and processing, data generation, validation, and analyses.

2.6. Acknowledgements

We thank M. Horn for helping with processing macaque tissue, Y. Imamura Kawasawa and M. Li for helping with human tissue and data analyses, and the members of the Sestan laboratory for valuable comments. Human post-mortem tissue was obtained from sources listed in Supplementary Information. Support for predoctoral fellowships was provided by the Portuguese Foundation for Science and Technology (A.M.M.S.), the Croatian Science Foundation (G.S.), the National Science Foundation (DGE-1122492; K.A.M.), and the China Scholarship Council (Y.Z.). This work was supported by grants from the NIH (MH081896, MH089929, NS051869), the Kavli Foundation and NARSAD, and by a James S. McDonnell Foundation Scholar Award (N.S.).

2.7. Figures

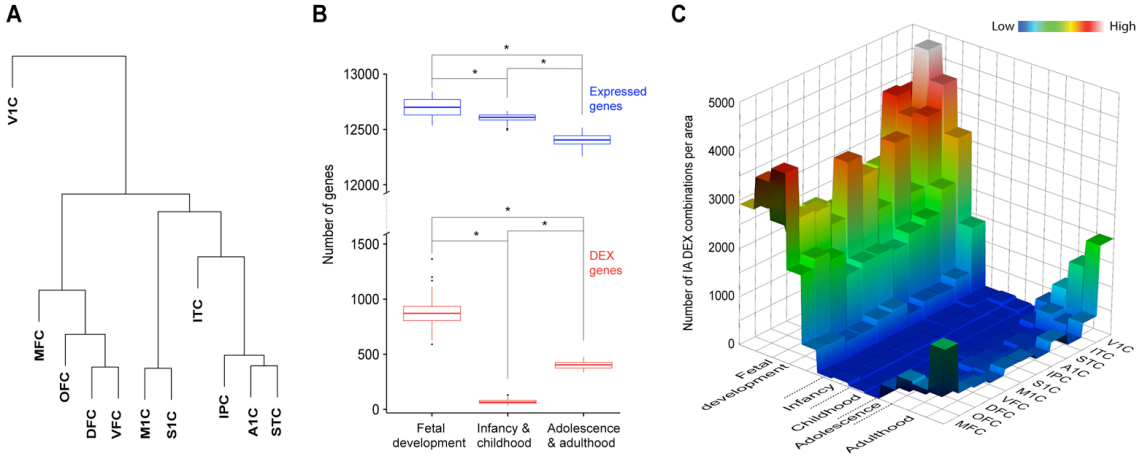


Fig. 1. Neocortical areal differences in gene expression exhibit a temporal hourglass pattern.

(**A**) Unsupervised hierarchical clustering of the 11 NCX areas profiled in this study, based on the transcriptome of each NCX area throughout the fetal development and postnatal lifespan, shows relative transcriptional differences. (**B**) Boxplots of subsampling permutations show the number of expressed (blue) and differentially expressed genes (red) among neocortical areas across fetal development (periods 3-7), infancy (periods 8 and 9) and childhood (periods 10 and 11), and adolescence (period 12) and adulthood (periods 13-15). (**C**) Post-hoc Tukey tests were used to identify significant comparison of inter-areal (IA) DEX combinations. A 3D heatmap shows the number of IA DEX combinations per NCX area across time.

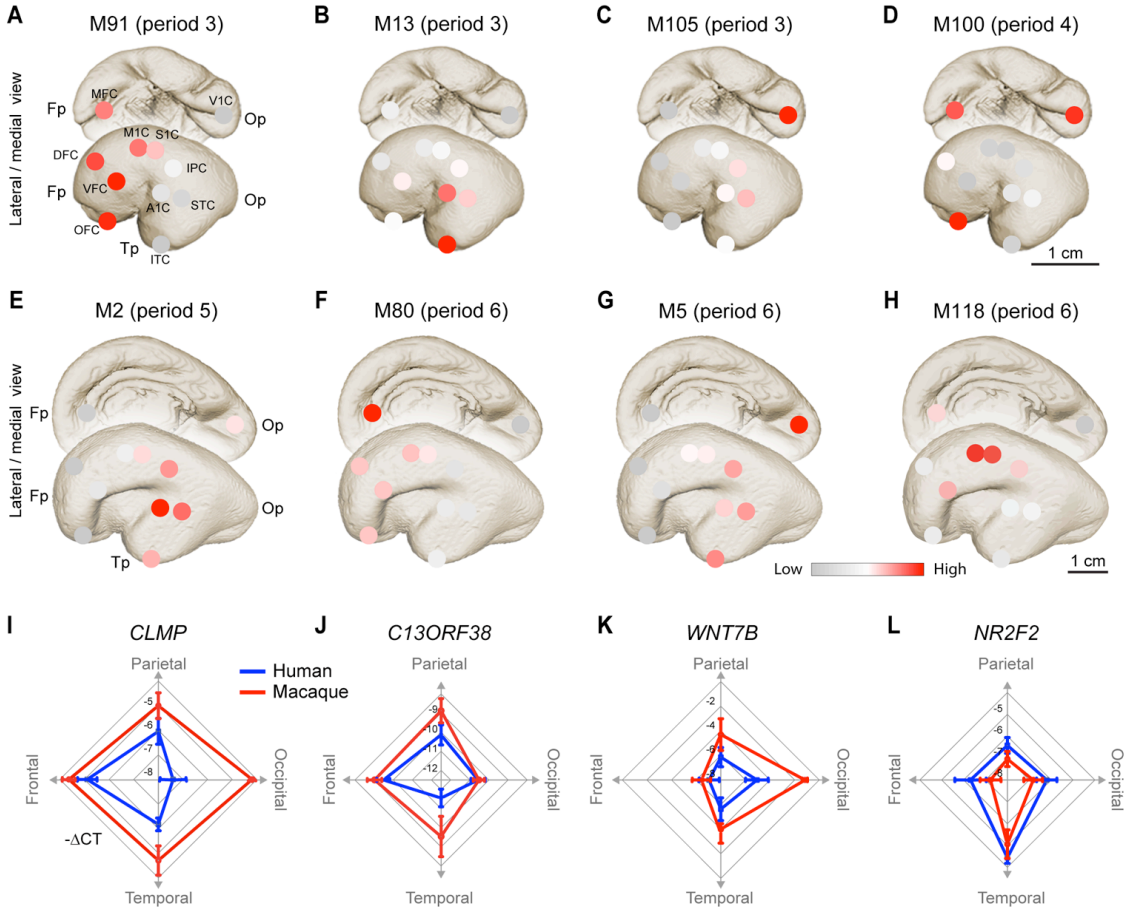


Fig. 2. Fetal gene co-expression modules illustrating the diversity of human spatial patterns and inter-primate similarities and differences.

(**A**-**H**) Average scaled-expression of all genes in co-expression module (M) 91 (**A**), M13 (**B**), M105 (**C**), M100 (**D**), M2 (**E**), M80 (**F**), M5 (**G**), and M118 (**H**) show a gradient-like expression pattern in periods 3 (**A**, **B**, **C**), 4 (**D**), 5 (**E**), and 6 (**F**, **G**, **H**). (**I**-**L**) Radar charts of qRT-PCR of intra-modular hub genes *CLMP* (M91) (I), *C13ORF38* (M80) (J), *WNT7B* (M6) (**K**), and *NF2R2* (M13) (**L**) demonstrate a gradient-like expression pattern in human (blue) in directions of areas grouped by their corresponding lobes (e.g., frontal, parietal, occipital and temporal; versions of these radar charts with all 11 areas across different periods are shown in figure S14). Rhesus macaque (red) expression patterns show similarities and differences to human.

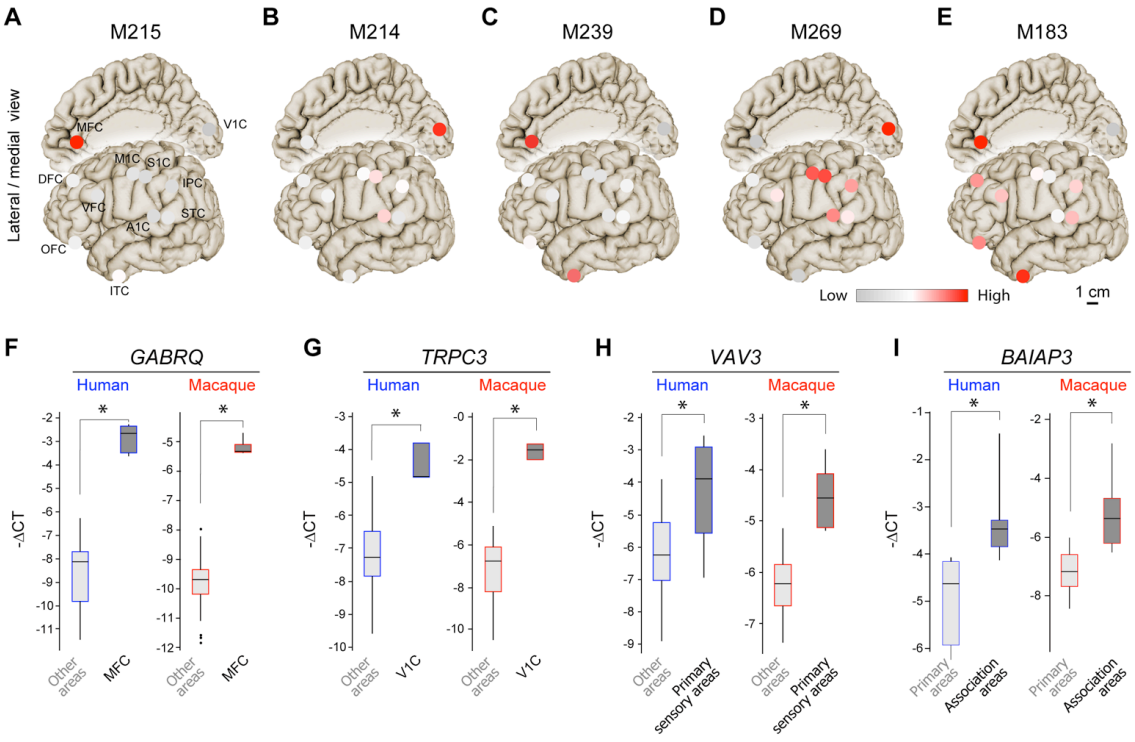


Fig. 3. Adolescent/adult gene co-expression modules illustrating the diversity of human spatial patterns and inter-primate similarities and differences.

(**A**-**E**) Average scaled-expression of all genes in M215 (**A**), M214 (**B**), M239 (**C**), M269 (**D**) and M183 (**E**), in period 13, show an area-specific (**A**, **B** and **C**) or a primary sensory areas-specific (**D** and **E**) patterns. (**F**-**I**) Box plots show qRT-PCR expression levels of area-enriched human (blue) and rhesus macaque genes (red): *GABRQ* (M215) (**F**), *TRPC3* (M214) (**G**). Genes with patterns that discriminate primary sensory areas and associative areas are also shown: *VAV3* (M214) (**H**) and *BAIAP3* (M183) (**I**), respectively. Radar charts of qRT-PCR data for the four genes with all 11 areas across different periods are shown in figure S16.

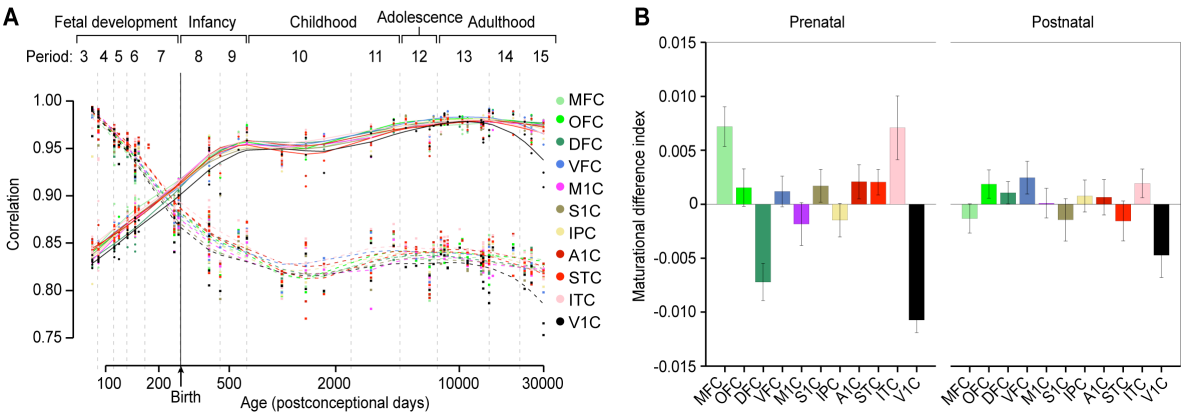


Fig. 4. Areal transcriptional trajectories become more synchronized during postnatal development.

(**A**) A maturational trajectory plot shows the Pearson correlation of gene expression in each sample to the corresponding averaged gene expression in period 14 (solid line) or period 3 (dashed line). (**B**) Bar plots show the average deviation of the areal trajectory from the average overall maturational trajectory (maturational difference index). Prenatal (fetal) development exhibits more deviation than postnatal development.

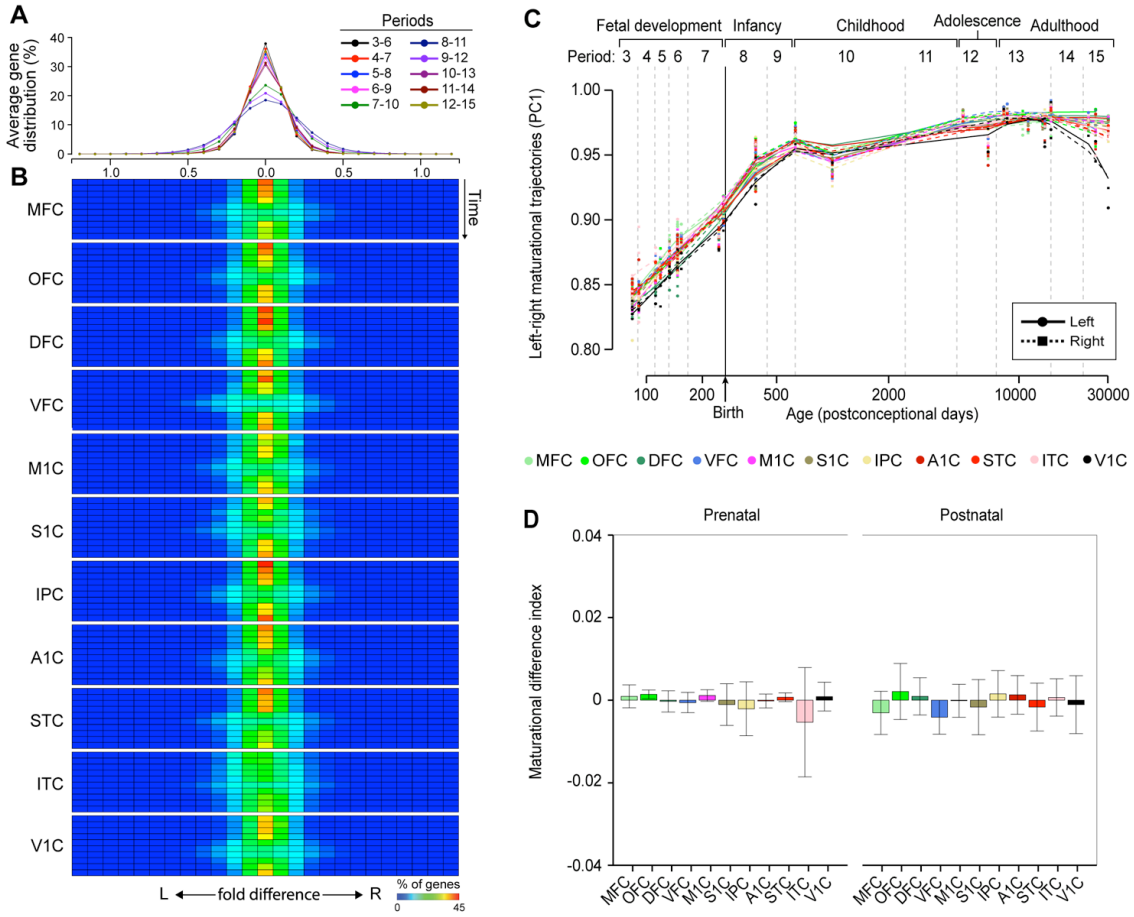


Fig. 5.Areal transcriptomes are globally symmetric across fetal development and postnatal lifespan.

(**A**) Average distribution of hemispheric bias for all profiled neocortical areas through overlapping time points of development and adulthood. (**B**) A heat map shows the distribution of fold differences in expression of all analyzed genes between left and right hemispheres for each profiled neocortical area across time. Each row corresponds to a 4-period sliding window (first row: 3-6; last row: 12-15). Maturational trajectories of left (solid line) and right (dashed line) hemispheres of each NCX area (**C**), and maturational difference index between hemispheres in prenatal (fetal) and postnatal development (**D**) further demonstrate a symmetric gene expression profile.

2.8. Supplementary Information

Table of Contents

1. Introduction

2. Neocortical Areas and Time Periods

3. Transcriptome Analyses
    - 3.1. Gene-level analysis
    - 3.2. Principal component analysis
    - 3.3. Identification of spatial DEX genes
    - 3.4. DEX analysis quality control
        - 3.4.1. Subsampling
        - 3.4.2. Permutation test
    - 3.5. Transcriptome maturation analyses
        - 3.5.1. Maturational trajectories
        - 3.5.2. Maturational difference index
    - 3.6. Creation of gene lists and transcriptional profiling of neurobiological categories
    - 3.7. Weighted gene co-expression network analysis (WGCNA)
        - 3.7.1. Dataset filtering
        - 3.7.2. Network construction and module detection
    - 3.8. Gene ontology (GO) enrichment analysis
    - 3.9. Spatio-temporal dynamics of co-expression modules
    - 3.10. Analysis of left- right bias in gene expression

4. Transcriptome Validation Methods
    - 4.1. Quantitative real-time RT-PCR

5. Dissection of Rhesus Macaque (*Macaca mulatta*) Neocortical Areas

6. Supplementary Tables
    - **Table S1 |** Periods of human brain development as defined in this study
    - **Table S2 |** Ontology and nomenclature of analysed NCX areas
    - **Table S3 |** Gene ontology of DEX genes

Table S4 | Fetal WGCNA modules

Table S5 | Adolescence and adulthood WGCNA modules

Table S6 | List of qRT-PCR primers

7. Supplementary Figures

Figure S1 | Principle component analysis

Figure S2 | PCA clustering of the samples with RIN as the confounding variable

Figure S3 | PCA clustering of the samples with PMI as the confounding variable

Figure S4 | PCA clustering of the samples with pH as the confounding variable

Figure S5 | Number of expressed and inter-areal DEX genes per period

Figure S6 | Trajectories of genes associated with neurodevelopmental processes (cell proliferation related genes, doublecortin, axon development related genes)

Figure S7 | Trajectories of genes associated with neurodevelopmental processes (dendrite development, synapse development and myelination related genes)

Figure S8 | Trajectories of genes associated with cortical glutamatergic neurons

Figure S9 | Trajectories of genes associated with cortical GABAergic interneurons

Figure S10 | Trajectories of genes associated with glial cells

Figure S11 | Trajectories of genes associated with neurotransmitter receptors (acetylcholine and serotonin receptor genes)

Figure S12 | Trajectories of genes associated with neurotransmitter receptors (glutamate and dopamine)

Figure S13 | Select prenatal WGCNA modules

Figure S14 | Prenatal qRT-PCR validations

Figure S15 | Select adolescent and adult WGCNA modules

Figure S16 | Adolescent and adult qRT-PCR validations

1. Introduction

In this Supplementary Information we provide further information regarding materials and methods. For a detailed description of the collection, dissection methods, and quality control assessments of post-mortem human brain, please refer to our previous study (*38*). Tissue was obtained from several sources including the Human Fetal Tissue Repository at the Albert Einstein College of Medicine, the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, the Laboratory of Developmental Biology at the University of Washington (supported by grant HD000836 from the Eunice Kennedy Shriver National Institute of Child Health and HumanDevelopment) and the JointMRC/Wellcome Trust HumanDevelopmental Biology Resource (http://hdbr.org) at the IHG, Newcastle Upon Tyne (UK funding awards G0700089 and GR082557). We also make available additional data that were discussed in the main manuscript. Finally, we present supplementary figures and tables generated from sample metadata and specific gene lists.

2. Neocortical Areas and Time Periods

For this study, we used a subset of a dataset previously reported *(38)*. In that dataset, the full course of human brain development and adulthood was divided into 15 periods. In the present study, we analysed 11 areas of the neocortex (NCX) in periods 3-15, an interval during which the areas can be consistently followed across time. The first 2 periods, which correspond to the embryonic and the first two weeks of the early fetal development, were not analysed in the present study since they don't contain the same 11 areas from periods 3-15.

The complete list of periods and NCX areas analysed in this study can be found in the Supplementary Tables 1 and 2, respectively.

3. Exon Array Data Analyses

3.1. Gene-level analysis

The expression levels for all core probe sets within an exon were averaged to obtain an expression value for the exon, and the median of all exons within one gene (transcript cluster) was used as the estimate of gene expression. To reduce the noise in the dataset and the false positives in the following differentially expressed genes investigations, we excluded genes with a log2-transformed expression value <6 in all samples. To consider a gene "expressed" in a particular area or at a particular time period, we required it to have a mean DABG P<0.01in that NCX area.

3.2. Principal component analysis

The whole data set was subjected to the principal component analysis (PCA). The first and the second principal components (PC1 and PC2) were plotted and samples were colored by periods (Figure S1). To explore the effects of confounders, we plotted the first four principal components against RIN (Figure S2), PMI (Figure S3) and pH (Figure S4), and colored samples by periods.

3.3. Identification of spatial DEX genes

Analysis of variance (ANOVA) was used to identify differentially expressed (DEX) genes across all regions of interest. Genes that were DEX in at least one NCX area were identified based on ANOVA by using area as an ANOVA factor. In order to exclude the possibility that variation in RIN scores and PMI within the acceptable range might influence procedures, we included PMI and RIN as technical covariates within our ANOVA model of differential expression. Resulting *P*-values from ANOVA were corrected for multiple comparisons using the Benjamini and Hochberg false discovery rate (FDR) method *(51)*. A conservative statistical threshold (FDR <0.01 and minimum fold difference >2 between NCX areas) was used to identify DEX genes. Genes that were not significantly expressed above background were excluded from ANOVA tests. A post-hoc Tukey's HSD (honestly significant difference) test was used to calculate differences between each pair of neocortical areas, in each period. The number of significant comparisons for each area, per period, was then used as the inter-areal (IA) DEX combinations number.

3.4. DEX analysis quality control

3.4.1. Subsampling

Random subsamples were used to calculate the number of expressed and DEX genes. All samples were divided into three groups: prenatal, postnatal, and adult. The number of samples randomly chosen was calculated as ¾ of the lowest number of samples contained within any of the three groups. For 100 trials, this fixed number of samples was randomly drawn from each group and the number of expressed or DEX genes was calculated using the selected subset of samples.

3.4.2. Permutation test

A permutation test was used to determine the dependence between group label and the mean number of expressed or DEX genes in the subsampling results. For each pair of groups (prenatal-postnatal, prenatal-adult, postnatal-adult), the test statistic was computed by subtracting the group with the smaller mean number from the group with the larger mean number. To calculate the null statistic, the group labels were shuffled, and

mean difference was calculated retaining the same group order used to calculate the test statistic. The permutation was performed 999,999 times to generate the null distribution and calculate a one-sided p-value.

3.5. Transcriptome maturation

3.5.1 Maturational trajectories

Period 14 is characterized by an adult-like pattern of brain functions (*38*), and can be considered as the period when cortical areas reach full maturation. We calculated the mean gene expression for each NCX area in this period and used it as the fully mature NCX reference for further comparisons. We then used the Pearson correlation of each sample to its areal reference as a measure of maturation. The values of correlation were then smoothed to fit a trajectory by R function smooth.spline.

3.5.2 Maturational difference index

To better understand differences among areas in maturation, we calculated the average maturational trajectory for all areas. Areal deviations from the average maturational trajectory were used to calculate the maturational difference index for each area, in each period. The average maturational difference index was then calculated for prenatal and postnatal development.

3.6. Creation of gene lists and transcriptional profiling of neurobiological categories

We manually curated lists of genes functionally related to specific neurodevelopmental processes, neural cell types, and neurotransmitter systems, by selecting individual genes or a small group of genes from the existing gene ontology (GO) database (http://www.geneontology.org/) whose functions, and thus expressions, have been most closely associated with specific neurobiological category based on published data from model organisms and human medical genetics. Detailed information on gene selection and neurobiological categories are provided in ref. *38* and *www.humanbraintranscriptome.org* database. We expected that the selected genes or gene group may accurately reflect the trajectories of neurobiological processes in neurodevelopment. To summarize the principle gene expression profile of each category, PCA was performed. The resulting first principal component (PC1) was plotted against logarithmic age in days, and a smooth curve was fitted by smoothing spline to display the developmental trajectories. For categories with only one gene, the expression level was used as PC1.

To compare the spatio-temporal patterns of genes in the same group, we plotted genes in each group individually and fitted the individual gene expression pattern by

smoothing spline. To compare the developmental trajectories for neurodevelopmental processes (Fig. S6 and S7), glutamatergic neurons (Fig. S8), interneurons (Fig. S9), glial cells (Fig. S10), and neurotransmitter receptors (Figs. S11 and S12), PC1 of each group was subjected to smoothing spline curve fitting. The fitted value $\hat{y}_i$ was then scaled to 0 to 1. The range of fitted vector $\hat{Y}$ is set to 1, and the scaled values were calculated as $z_i = (\hat{y}_i - \min(\hat{Y}))/\text{range}(\hat{Y})$. The scaled values were then plotted to represent the trajectory of each group with a measure of the percentage of changes. Furthermore, the average trajectory was calculated as a function of average value for all areas per period. Deviations from the average maturational trajectory were used to calculate the maturational difference index for all areas. The average maturational difference index was then calculated for prenatal and postnatal periods (3-7 and 8-15, respectively).

3.7. Weighted gene co-expression network analysis (WGCNA)

3.7.1. Dataset filtering

WGCNA was applied to samples from periods 3-6 and periods 12-15, respectively. Since the analysis is sensitive to noise, only genes with log2-expression value greater than 6 in any of the sample and coefficient of variation (CV) greater than 0.05 were used for analysis.

3.7.2. Network construction and module detection

Signed co-expression networks were built using the WGCNA package (*52, 53*) in R. For all genes included, a pair-wise correlation matrix was computed, and an adjacency matrix was calculated by raising the correlation matrix to a power. The power was set to 18 for both of Periods 3-6 and Periods 12-15, according to a scale-free topology criterion (*54*). For each pair of genes, a robust measure of network interconnectedness (topological overlap measure) was calculated based on the adjacency matrix. The topological overlap based dissimilarity was then used as input for average linkage hierarchical clustering. Modules were generated by hybrid dynamic tree-cutting. To obtain rare expression patterns, we set the minimum module size to 5 genes, deepSplit to 3, and the minimum height for merging modules to 0.15.

Each module was summarized by an eigengene, which is the first principal component of the scaled module expression. Thus, the module eigengene explained the maximum amount of variation of the module expression levels.

3.8. Gene ontology (GO) enrichment analysis

Functional enrichment was assessed using DAVID Bioinformatics Resources 6.7 *(55, 56)* (http://david.abcc.ncifcrf.gov/).

3.9. Spatio-temporal dynamics of co-expression modules

The expression patterns of modules were manually categorized based on module eigengenes. Modules that showed gradients in expression (Fig. S13) or region-specific expression (Fig. S15) were selected. To display the changes of expression across periods, the expression of genes in each selected module was scaled and the module expression was calculated by averaging the scaled expression across genes. The module expression was then plotted as a heatmap to show the change of gradients across periods. To demonstrate gradients for a selected period, the module expression in that period were plotted on a reference brain and labeled by gradients of colors.

3.10. Analysis of left-right bias in gene expression

Differences in gene expression between left and right hemispheres were analyzed by the combination of a sliding window algorithm, and a paired t-test. The sliding-window algorithm was used to detect left/right gene expression differences within a group of sequential periods. The window size was set to span 4 periods. For each window, a paired t-test was applied to determine if the expression level of a gene in left hemispheres was significantly different from the expression level in right hemispheres in the same area. The *P*-values from the t-tests were transformed to FDR using the Benjamini and Hochberg method *(51).* For each gene, the fold difference (log2-transformed) between left and right hemisphere samples in each region was also calculated. An FDR of 0.1 was used as a cutoff to identify genes that are DEX between left and right hemispheres in each window.

Using the same approach as in 3.5., we calculated the maturation trajectories of both hemispheres. The mean areal correlation of each period was calculated by hemisphere. The difference between hemispheres was then calculated for each area in each period. Bar plots were used to summarize the difference between hemispheres in prenatal and postnatal development.

4. Transcriptome Validation Methods

4.1. Quantitative real-time RT-PCR

An aliquot of the total RNA that was previously extracted from each brain region was used for secondary validation through real-time PCR analysis. One µg of total RNA was used for cDNA synthesis using SuperScript III First-strand synthesis Supermix (Invitrogen) and subsequently diluted with nuclease-free water to 1 ng/µl cDNA. Gene-specific high-

melting temperature primers for genes of interest were designed using NCBI/Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and expressed sequence information obtained from GenBank (NCBI). PCR reactions were conducted on an ABI 7900 Sequence Detection System (Applied Biosystems) using a hot start SYBR-green based method (Fast SYBR Green Master Mix, ABI) followed by melt curve analysis to verify specificity of the product. The Ct value (cycle number at threshold) was used to calculate the relative amount of mRNA molecules. The Ct value of each target gene was normalized by subtraction of the Ct value from housekeeping genes to obtain the ΔCt value. The relative gene expression level was shown as -ΔCt. All genes of interest were normalized to the housekeeping gene *RPL32*. Also, in Table S6 we provide sequences of primers used in validation.

qRT-PCR results were shown in radar charts, which were plotted using Origin 8.6 plotting software. Gradient-like patterns of gene expression (Fig. 2I, J, K, L) were shown by grouping areas per lobe: frontal (MFC, OFC, DFC, VFC, M1C), parietal (S1C, IPC), temporal (STC, A1C, ITC), and occipital (V1C). Region specific/enriched patterns of genes (Fig. 3F, G, H, I) were plotted by grouping regions of interest against all other areas combined. Also, we show the expression values of each area using radar charts (Fig. S14 and S16) and we compared patterns of expression between the species, calculating Pearson's correlation coefficient for brains with matched age.

5. Dissection of Rhesus Macaque (*Macaca mulatta*) Neocortical Areas

We performed qRT-PCR using RNA isolated from homologous areas of the rhesus macaque (*Macaca mulatta*) NCX at equivalent developmental periods using a mathematical model to translate neurodevelopmental time across species *(57)* (www.translatingtime.net). Macaque experiments were carried out in accordance with a protocol approved by Yale University's Committee on Animal Research and NIH guidelines. All brains were dissected using the same protocol used for human specimens. Areas were sampled according to ref. *(58)*.

6. Supplementary Tables

Supplementary Tables are provided in a single Microsoft Excel file on the CD rom.

Table S1. Periods of human brain development as defined in this study

List of the periods of human brain development used in this analysis and their corresponding age.

Table S2. Ontology and nomenclature of analysed NCX areas

List of the neocortical areas analyzed and their corresponding Brodmann areas.

Table S3. Gene Ontology of DEX genes

List of gene ontology analysis performed by DAVID algorithm for DEX genes during fetal development, infancy and childhood, and adolescence and adulthood; list of fetal-only, and adolescence- and adulthood-only terms.

Table S4. Pre-natal (periods 3 to 6) co-expression modules 1-122

List of 122 modules corresponding to spatio-temporal patterns identified by weighted gene co-expression network analysis (WGCNA) with its functional interpretation and result of gene ontology analysis performed by DAVID algorithm; list of intra-modular genes and their intra-modular connectivity; and list of modules with gradient-like patterns: anterior-posterior, temporal, posterior perisylvian, and occipital.

Table S5. Post-natal (periods 12-15) co-expression modules 123-329

List of 206 modules corresponding to spatio-temporal patterns identified by WGCNA with its functional interpretation and result of gene ontology analysis performed by DAVID algorithm; list of intra-modular genes and their intra-modular connectivity; and list of modules with region-specific enrichment in V1C and MFC.

Table S6. List of primers used for qRT-PCR

List of human and macaque monkey primers used in this study. *RPL32* was used as the housekeeping gene.

7. Supplementary Figures

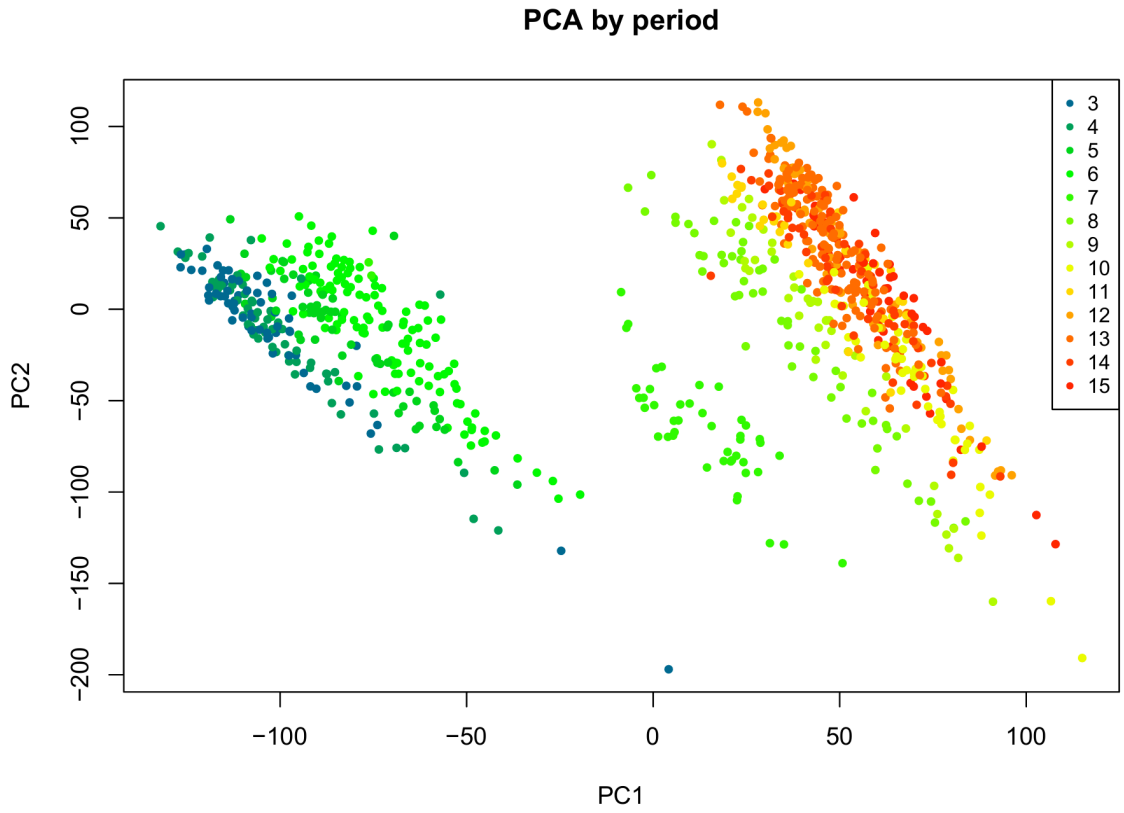


Figure S1 | Principle component analysis.

Principle component analysis, with samples colored by period, shows separation in three distinct clusters: early fetal to late mid-fetal (periods 3-6), late fetal (period 7) and postnatal (periods 8-15).

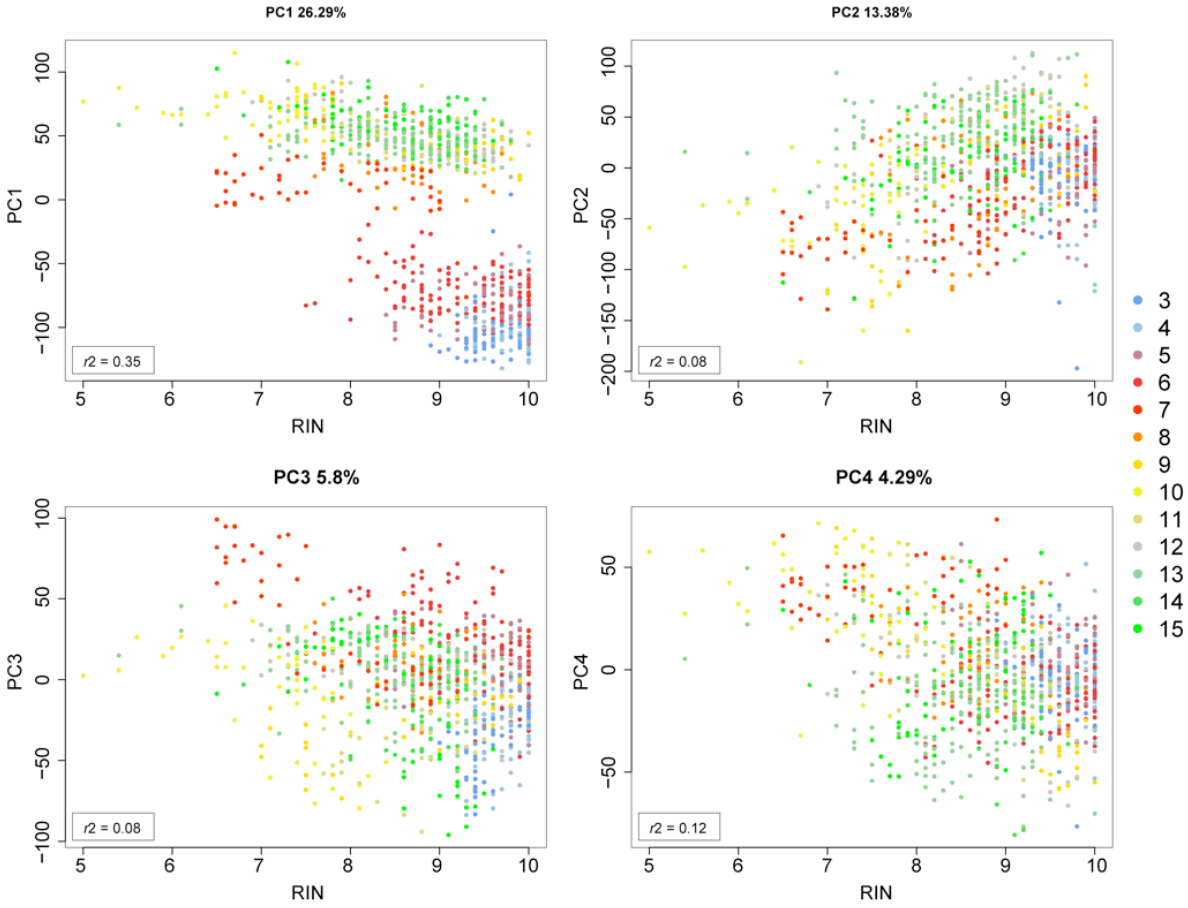


Figure S2 | PCA clustering of the samples with RIN as the confounding variable.
The first four principal components were plotted against RNA Index Number (RIN). Samples were colored by periods. Negligible correlation with the confounder was observed.

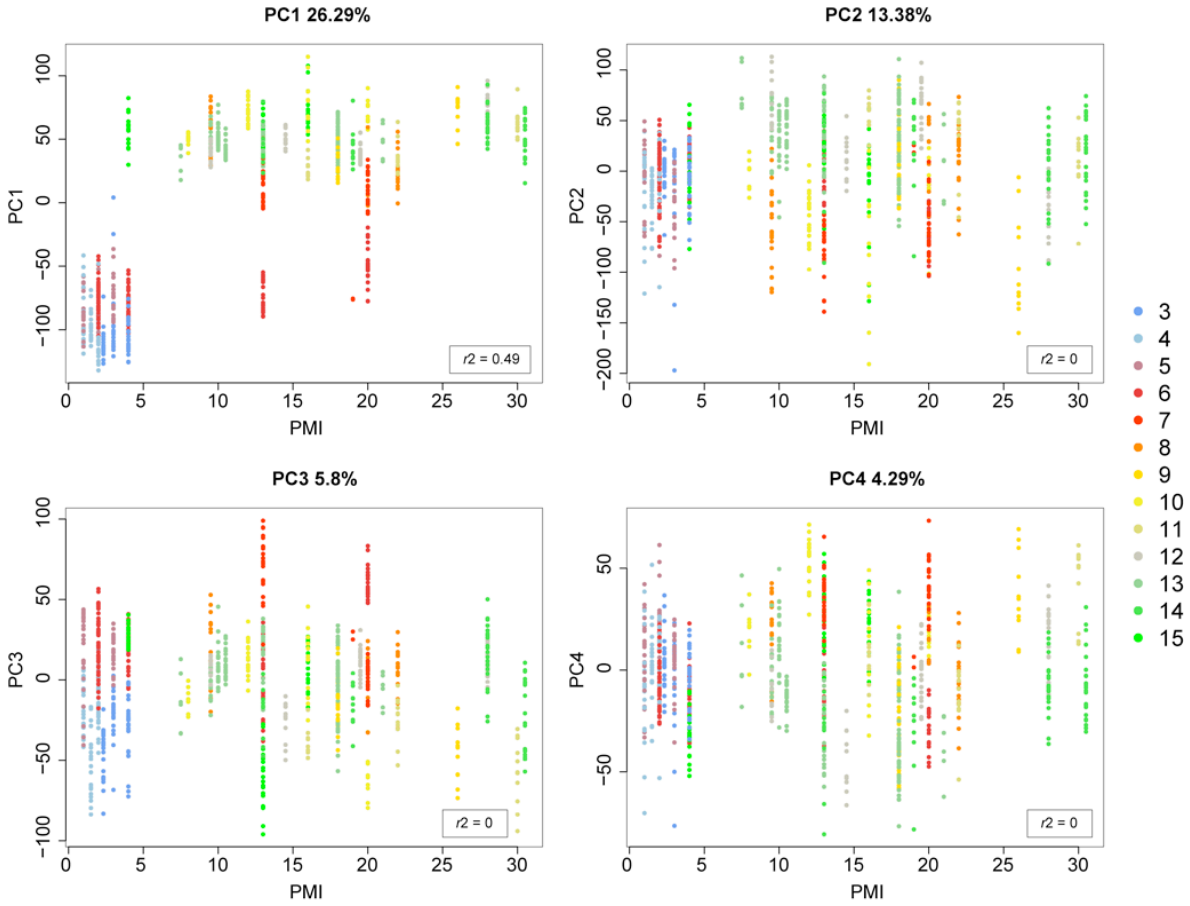


Figure S3 | PCA clustering of the samples PMI as the confounding variable.

The first four principal components were plotted against PMI. Samples were colored by periods. Negligible correlation with the confounder was observed.

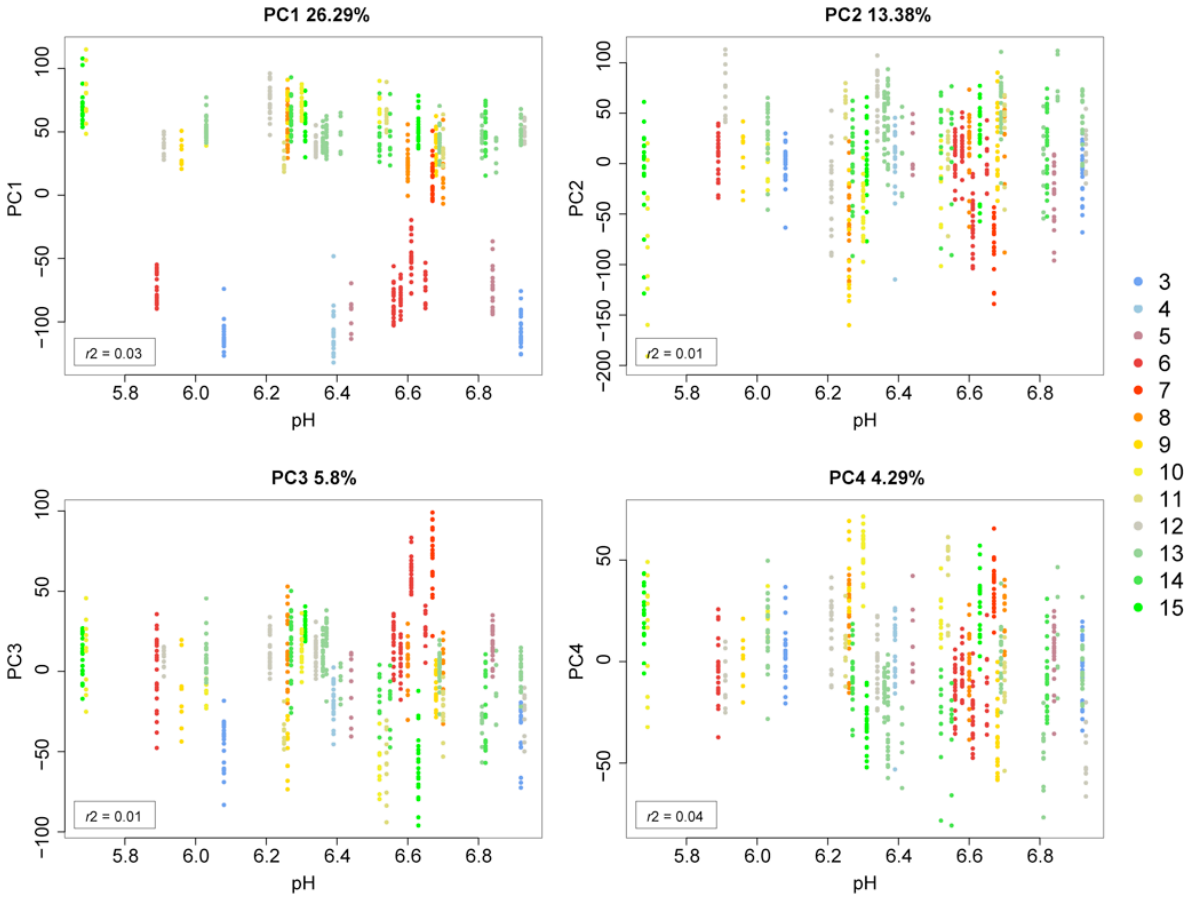


Figure S4 | PCA clustering of the samples pH as the confounding variable.

The first four principal components were plotted against pH. Samples were colored by periods. Negligible correlation with the confounder was observed.

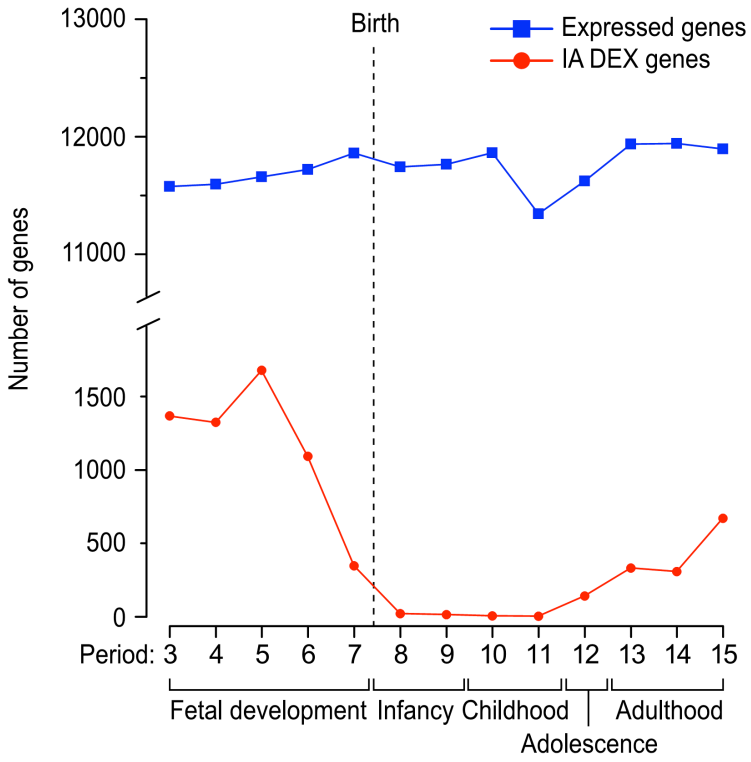


Figure S5 | Number of expressed and inter-areal (IA) DEX genes per period.
Analysis of DEX genes (red line) per period shows the hourglass shaped pattern, with the highest number of DEX genes during prenatal period, a lower number of DEX during infancy and childhood, and a reappearance of DEX from adolescence on. Number of expressed genes (blue line) stays stable throughout the fetal development and postnatal lifespan.

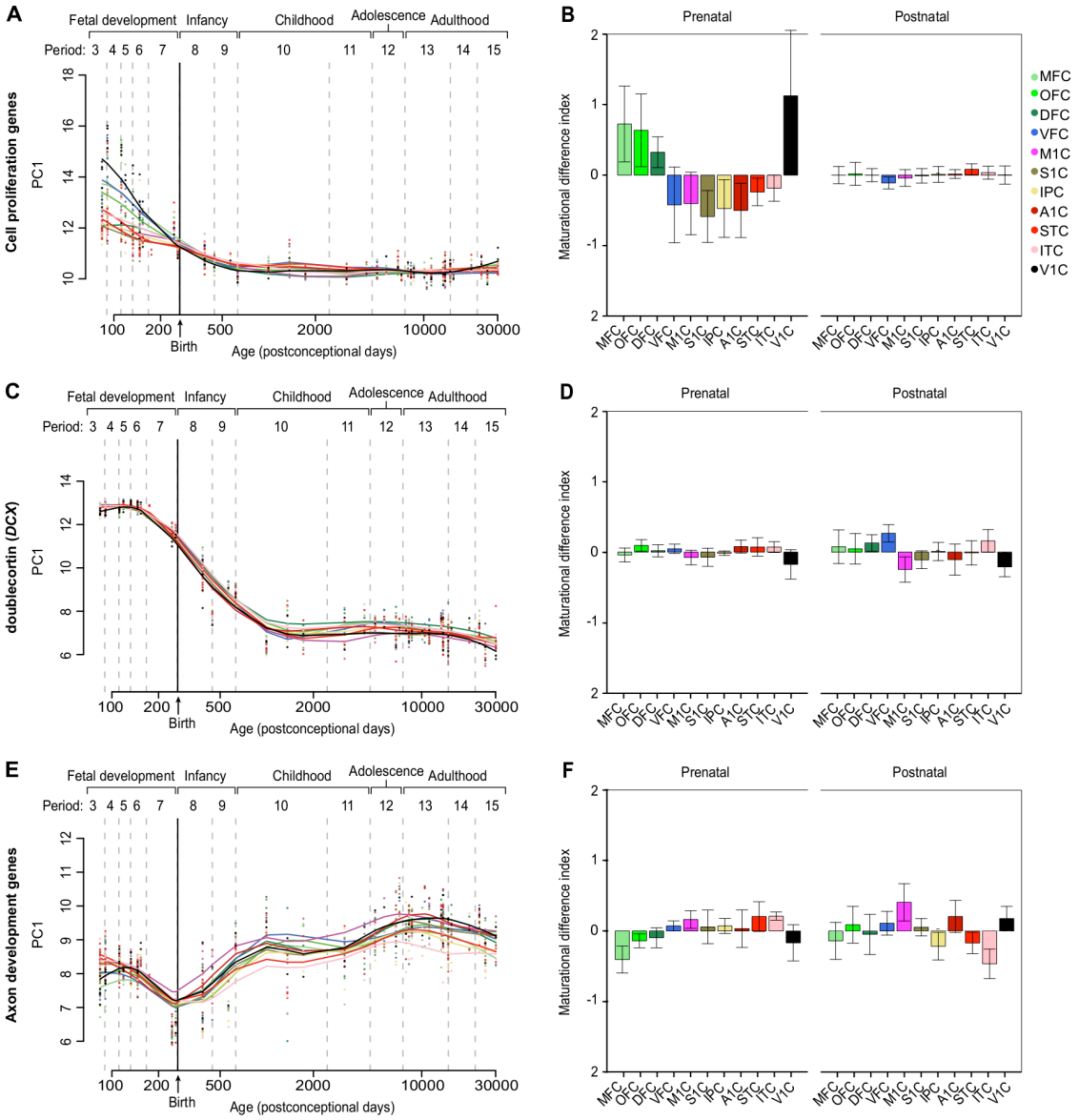


Figure S6 | Trajectories of genes associated with neurodevelopmental processes (cell proliferation related genes, doublecortin, axon development related genes). Genes associated with cell proliferation markers (*RRM1, HES1, NES*, and *MKI67*) are showing higher expression levels **(A)** and higher variability **(B)** during prenatal periods than after the birth. *Doublecortin* (*DCX*), a gene previously associated with neuronal migration and immature neuronal states, shows high expression **(C)** values during prenatal periods and a big decrease in expression levels during postnatal periods with little deviation from average trajectory **(D)**. Axon development marker (*CNTN2*) is showing triphasic trajectory **(E)** with higher expression during early and mid-fetal and postnatal periods but with decrease in expression levels around the birth. Deviation from average trajectory is present but relatively low in both prenatal and postnatal periods **(F)**.

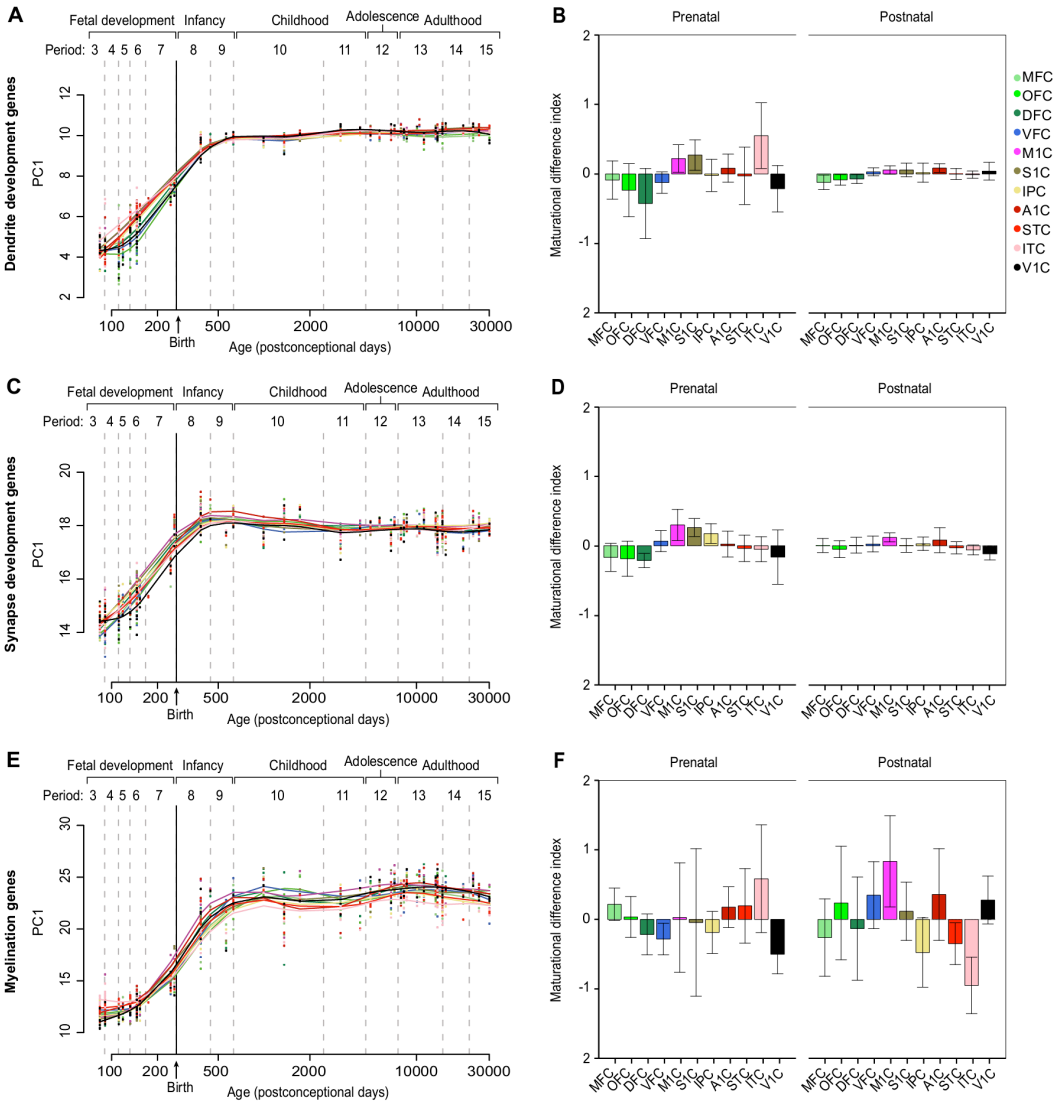


Figure S7 | Trajectories of genes associated with neurodevelopmental processes (dendrite development, synapse development and myelination related genes).

Dendrite development genes (*MAP1A*, *MAPT*, and *CAMK2A*) show gradual increases in expression until they reach the plateau at period 10 (**A**). At prenatal periods, areas show higher deviation from average trajectory than postnatal periods (**B**). Synapse development markers (*SYP*, *SYPL1*, *SYPL2*, and *SYN1*) show increased expression until postnatal periods when they reach a plateau (**C**). Deviation from the average trajectory, while present during prenatal periods, virtually disappears after birth (**D**). Myelination genes (*MBP*, *PLP1*, *MOG*, *C11orf9*, and *MAG*) show similar trajectories to previous the two processes with steady increase during prenatal periods and reaching the plateau in periods 9 and 10 (**E**), but deviation from the average trajectory is present in both prenatal periods (**F**).

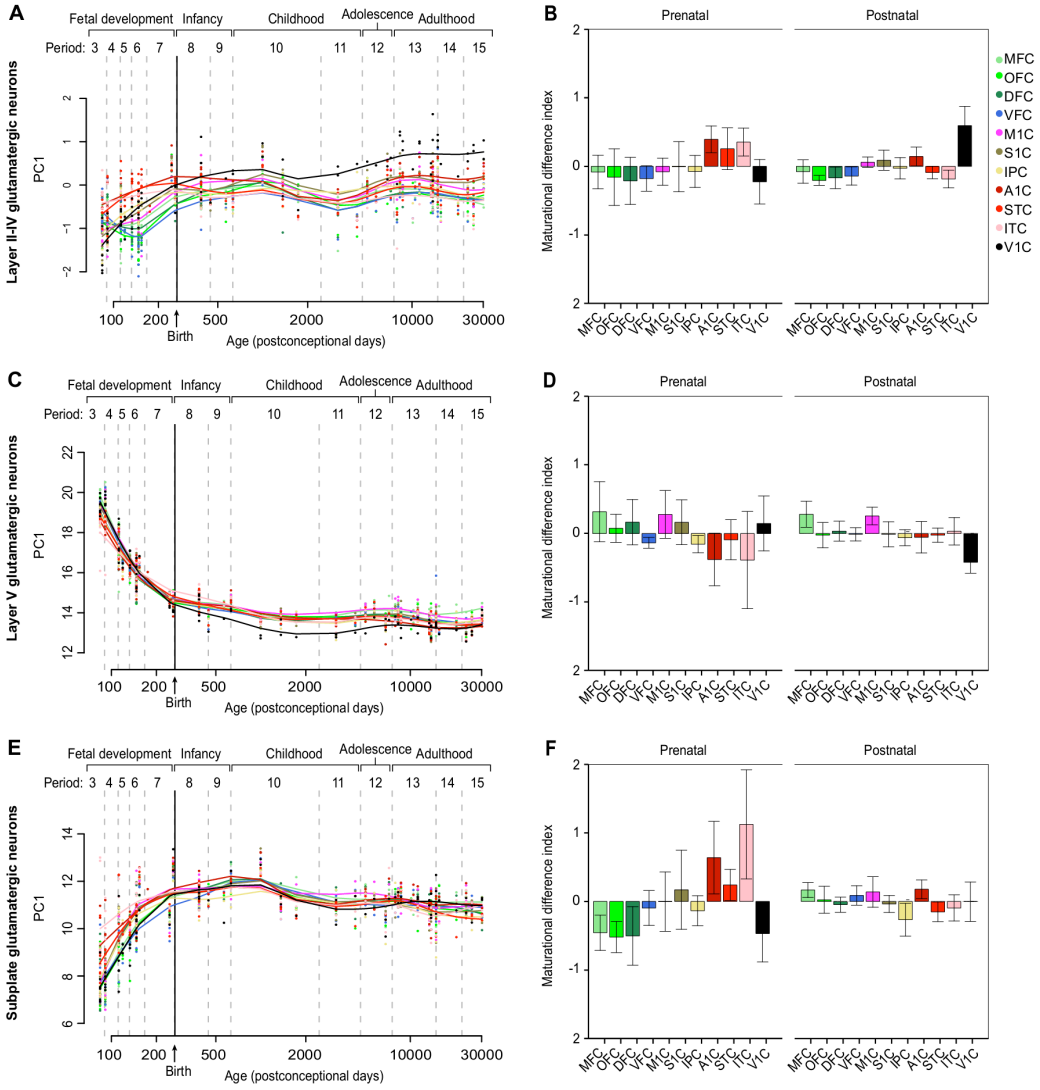


Figure S8 | Trajectories of genes associated with cortical glutamatergic neurons. Genetic makers for layer II-IV glutamatergic neurons (*CUX1* and *UNC5D*) show slow increases in expression levels during development (**A**). Exceptions to this are the frontal areas where in early fetal periods expression is higher than in late mid-fetal period. Variability from the average trajectory is present in both prenatal and postnatal periods, with highest deviation in postnatal V1C (**B**). Markers for layer V glutamatergic neurons (*FEZF2*, *BCL11B*, *OTX1*, and *ETV1*) show an opposite trajectory with the highest expression levels during prenatal periods and reach the lowest point during childhood (**C**). Deviations among the areas are larger during prenatal periods (**D**). Markers of subplate glutamatergic neurons (*CTGF* and *UNC5C*) show increased expression during prenatal periods reaching a maximum after birth and a gradual decrease during postnatal periods (**E**). Variation from average trajectory is larger during prenatal periods compared to postnatal periods (**F**).

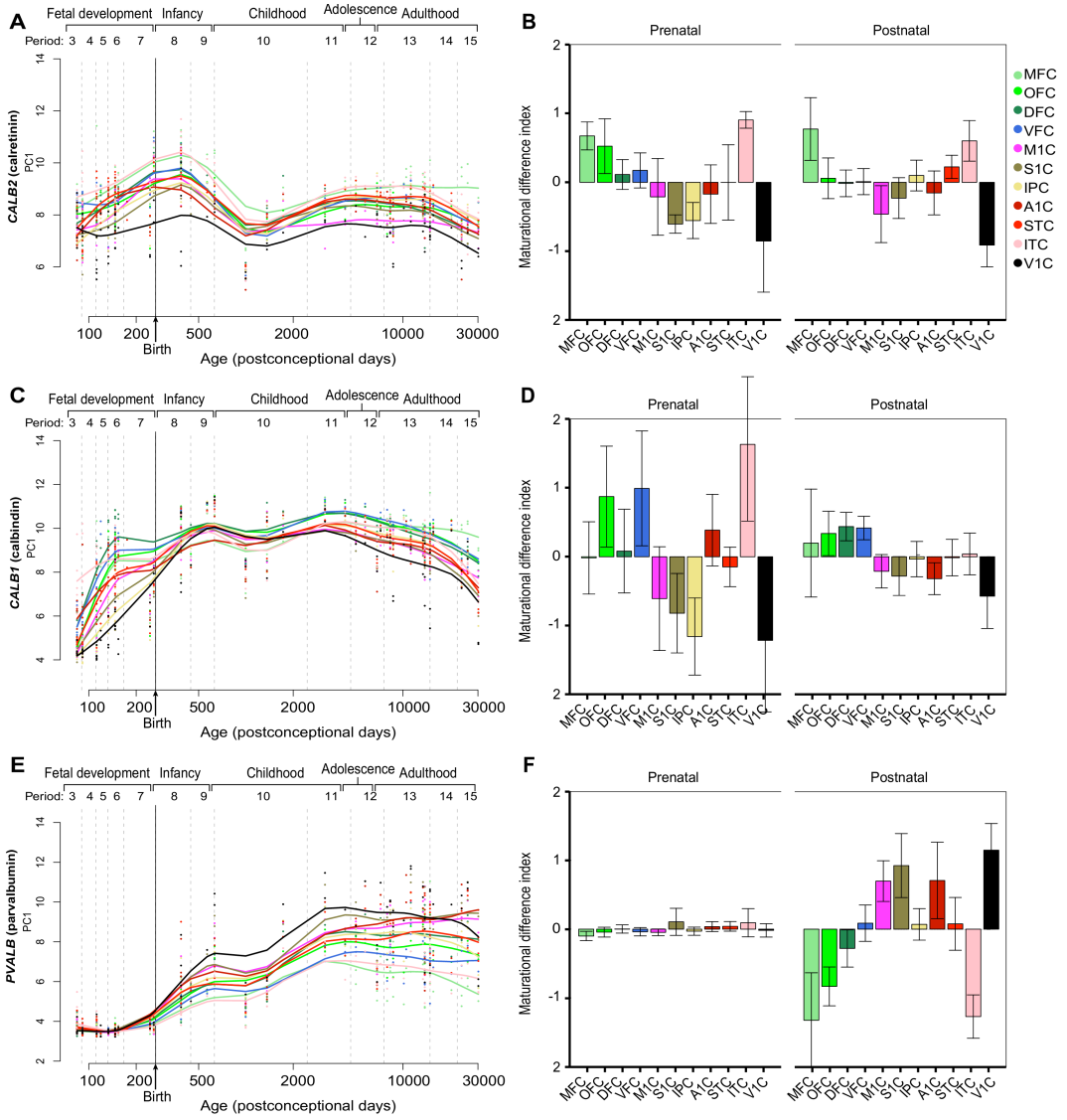


Figure S9 | Trajectories of genes associated with cortical GABAergic interneurons.

Calretinin (*CALB2*) shows increased expression until birth, when expression levels reach their maximum (**A**). After birth, expression levels start to decrease until period 10 when they reach their lowest values and then are followed by a small increase. All areas have high deviation from the average trajectory (**B**). Calbindin (*CALB1*) shows a highly dynamic trajectory, with an increase in expression during prenatal periods, with frontal regions increasing faster compared to others, especially V1C (**C**). Deviation from the average trajectory is present in both prenatal and postnatal periods, but much more accentuated prenatally (**D**). Parvalbumin (*PVALB*) expression levels gradually increase and reach their maximum at period 11 (**E**). Prenatal expression is low, while during postnatal periods expression levels are high with large deviations from the average trajectory (**F**).

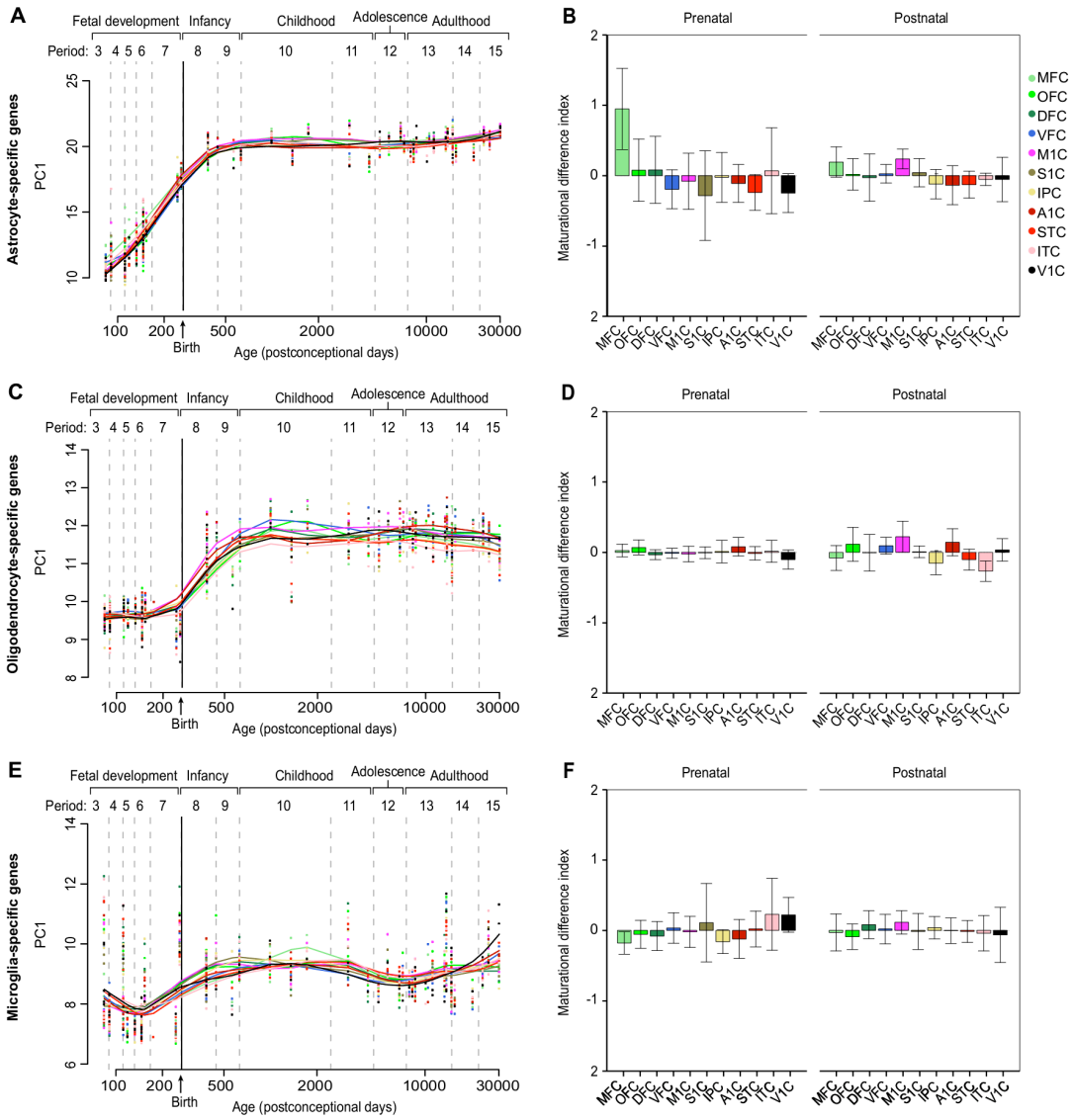


Figure S10 | Trajectories of genes associated with glial cells.

Astrocyte associated genes (*GFAP*, *S100B*, and *ALDOC*) gradually increase in expression and reach a plateau after birth (**A**). Prenatal MFC shows the highest deviation from average the trajectory (**B**). A mature oligodendrocyte genetic marker (*CNP*) shows a similar trajectory with lower deviation from the average trajectory prenatally (**C**, **D**). Microglial genes (*CFH*, *FCER1G*, and *TNIP2*) show less pronounced increased expression through development (**E**, **F**).

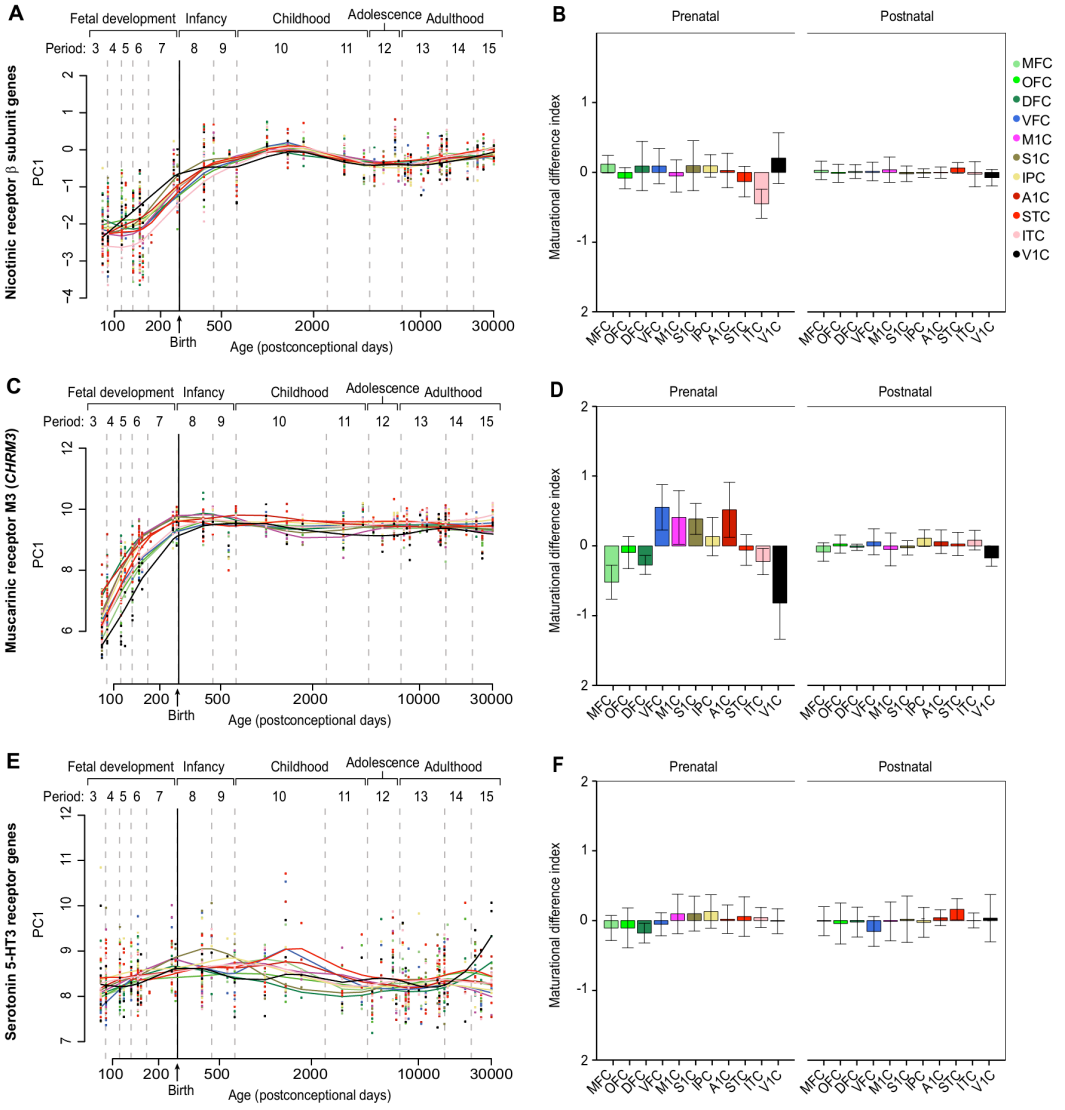


Figure S11 | Trajectories of genes associated with neurotransmitter receptors (acetylcholine and serotonin receptor genes).

Nicotinic receptor β subunit genes (*CHRNB1*, *CHRNB2*, *CHRNB3*, and *CHRNB4*) show a progressive increase in expression levels (**A**) with higher deviation from average trajectory during prenatal periods (**B**). Muscarinic M3 receptor gene (*CHRM3*) shows a sharper increase in expression levels prenatally reaching the maximum expression levels after birth (**C**). The deviation from average trajectory is higher during prenatal periods (**D**). Serotonin 5-HT3 receptor genes (*HTR3A*, *HTR3B*, *HTR3C*, *HTR3D*, and *HTR3E*) show expression levels that do not change drastically during the lifespan (**E**, **F**).

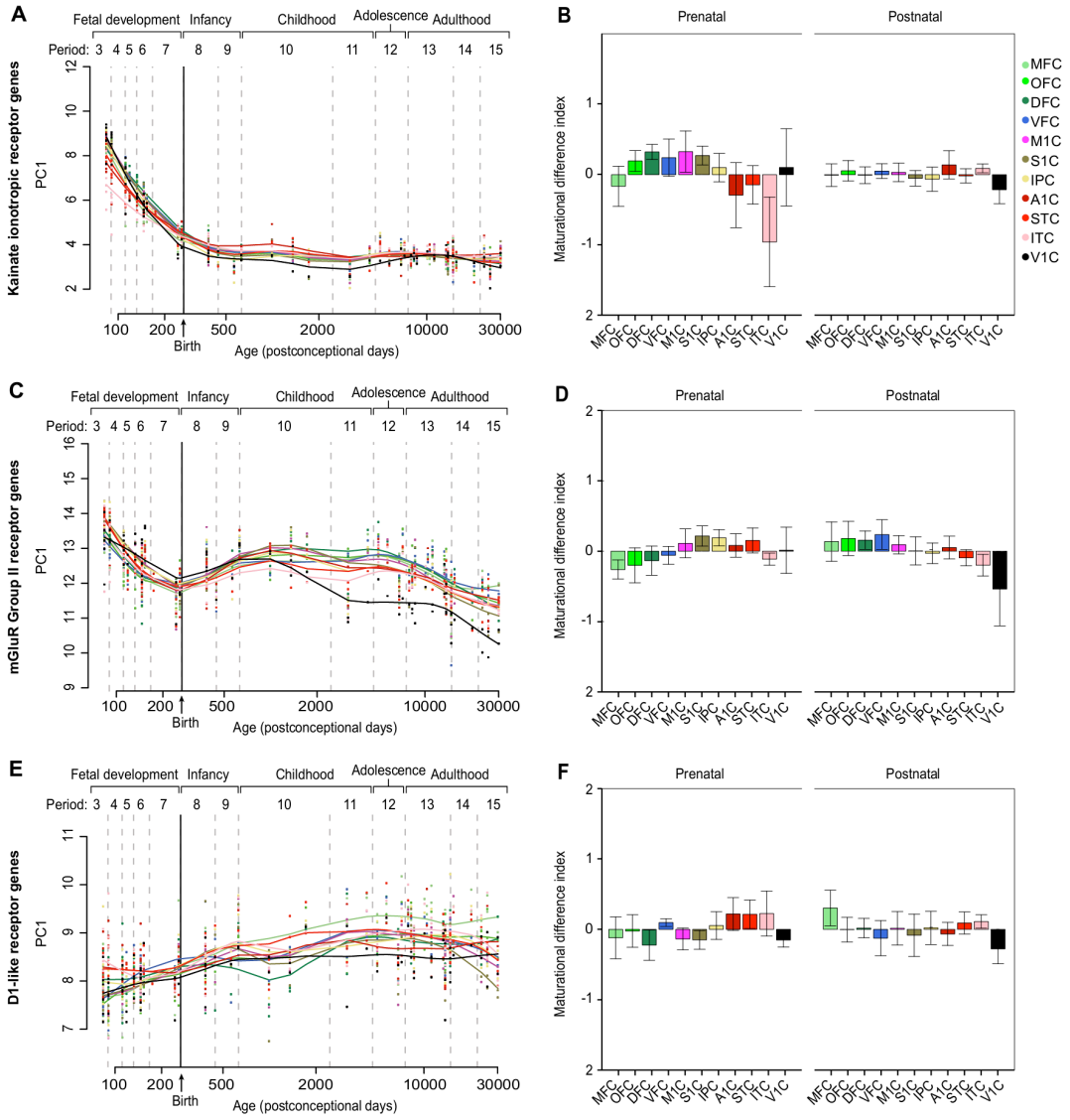


Figure S12 | Trajectories of genes associated with neurotransmitter receptors (glutamate and dopamine).

Kainate receptor genes (*GRIK1*, *GRIK2*, *GRIK3*, *GRIK4*, and *GRIK5*) show a sharp decrease in expression levels during prenatal periods (**A**) with higher deviation from average trajectory also prenatally (**B**). mGluR group 2 receptor genes (*GRM2* and *GRM3*) show a slower decrease in expression levels (**C**), with V1C in postnatal periods having the largest decrease in expression levels and the highest deviation from the average trajectory (**D**). D1-like receptor genes (*DRD1* and *DRD5*) show a slow increase of expression through the whole lifespan (**E**) with small but present deviations from the average trajectory (**F**).

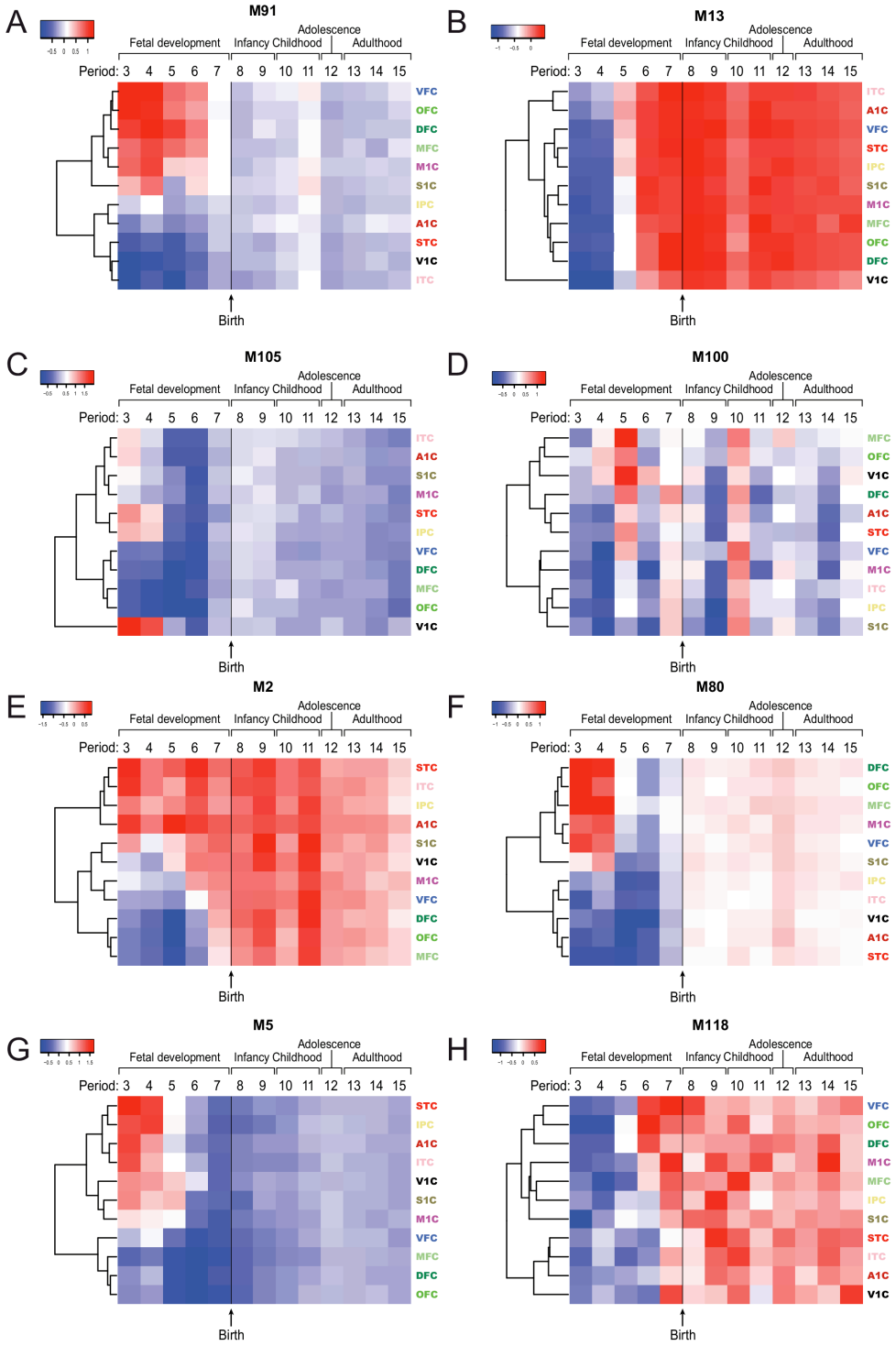


Figure S13 | Select prenatal WGCNA modules.

(**A**) Heat map showing module 91 (M91) with a frontal-cortex enriched gradient in periods 3-6 (top hub genes: *CYP26A1, HSPA12A, CNIH3, C8orf34, PDE4D, MANBA, ASAM, TSEN15, DPP7,* and *HS6ST3*). (**B**) M13 shows a temporal lobe enriched gradient in

period 5 (top hub genes: *LRRC4C, GABRB1, SLC24A3, KLHL4, SLC2A13, RIT2, GRM7, CSPG5, SV2B,* and *GDA*).(**C**) M105 contains genes with expression enriched in occipito-temporal areas in periods 3 and 4 (top hub genes: *SHISA6, KCNK12, DOCK10, PCK2, TMPRSS13, MAS1, STK10,* and *PPARG*), (**D**) M100 genes exhibit expression with a medio-lateral gradient in periods 4 and 5 (top hub genes: *APPL2, CTDSP2, HDAC1, ZNF43, NECAP2,* and *HEATR5*A), (**E**) M2 genes are enriched with posterior perisylvian and temporal expression levels in periods 3-5 (top hub genes: *PCDH9, CAMK4, OSTN, NCAM2, MDGA2, PLEKHH2,* and *COX7A1*), (**F**) M80 genes exhibit expression with frontally enriched expression levels (top hub genes: *YPEL2, DPF3, ZNF385A, GADL1, SLIT3, C13orf38, GNG7, TCERG1L, RCAN3,* and *FAM5B*), (**G**) M5 genes exhibit co-expression patterns in temporo-parieto-occipital areas in periods 3- 5 (top hub genes*: STARD13, PAPPA2, VRK2, SORCS1, NXPH3, KLHL32, CABP7, WNT7B, OCA2,* and *GPR126*), and (**H**) M118 genes exhibit expression highest in frontal areas in period 6 (top hub genes: *GDF10, ACHE, CYB561, PDE4A, ADAMTS8,* and *TMEM151A*).

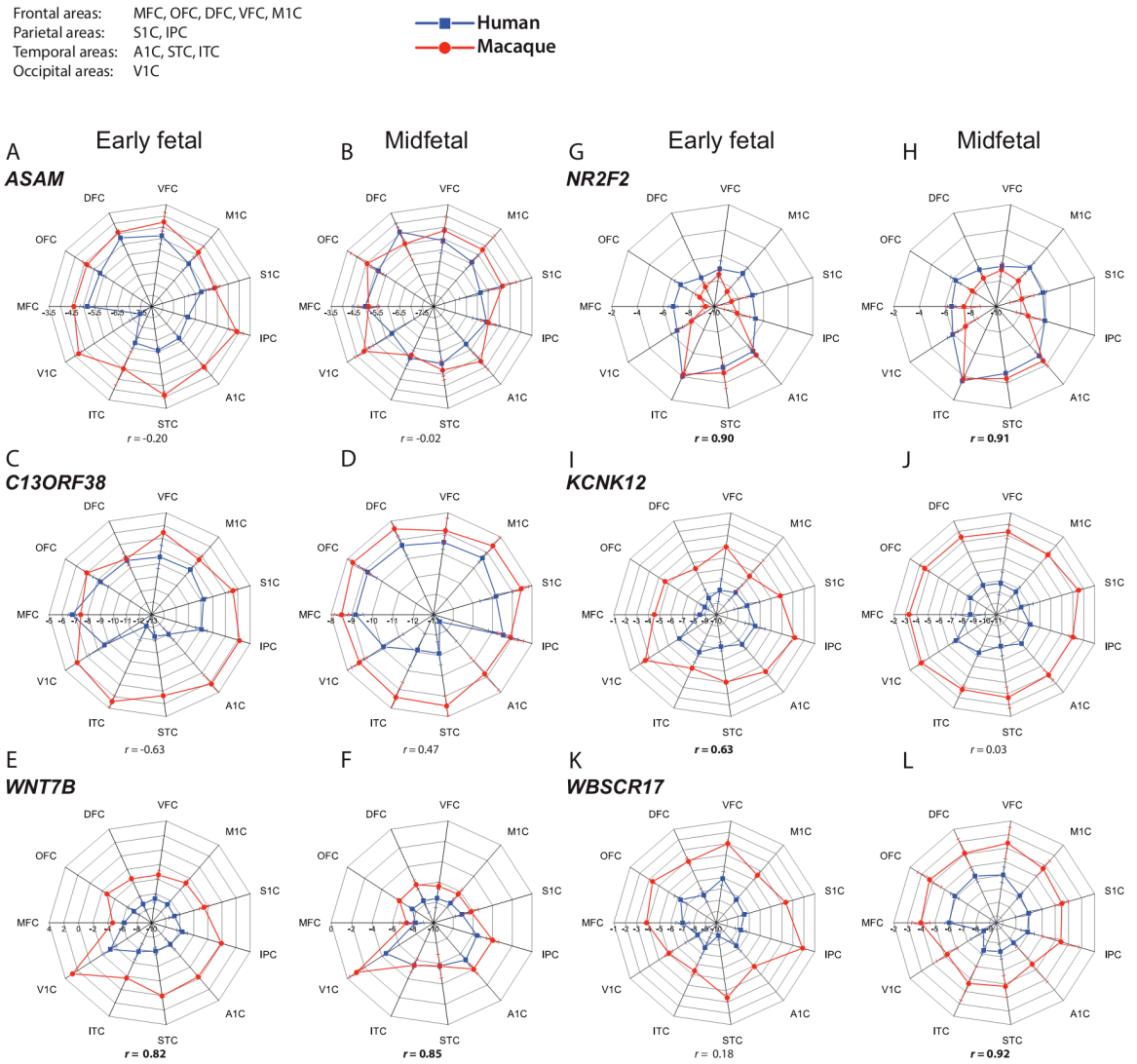


Figure S14 | Prenatal qRT-PCR validation.

Radar charts showing expression levels in all areas in prenatal periods 3-4 (early) and 5-6 (mid). Pearson's correlation coefficient *r* is printed bold when there is positive correlation between human and monkey patterns of expression. The expression pattern is shown with lines and points in blue color for humans and red color for macaque. *ASAM* **(A, B)** shows a different pattern of expression in human and in macaque telencephalon. In human cortex, *ASAM* shows a frontooccipital gradient during early fetal periods **(A)** which disappears in midfetal periods **(B)**. *C13orf38* **(C, D)** shows a fronto-temporal gradient in human braun in both early and midfetal periods (**C, D**, respectively), while in macaque NCX the fronto-temporal gradient is not existent; moreover, it is anti-correlated in the early fetal period **(C)**. *WNT7B* **(E, F)** shows the same gradient in human and monkey brains and in both early and mid-fetal periods. *NR2F2* **(G, H)** shows a temporal gradient in both species and in all prenatal periods. *KCNK12* **(I, J)** shows an occipito-frontal gradient in both periods in humans, while in macaque it is only during the early fetal period.

WBSCR17 **(K, L)** shows a correlated fronto-occipital gradient in both species during mid-fetal periods, which is not evident during the earlier periods.

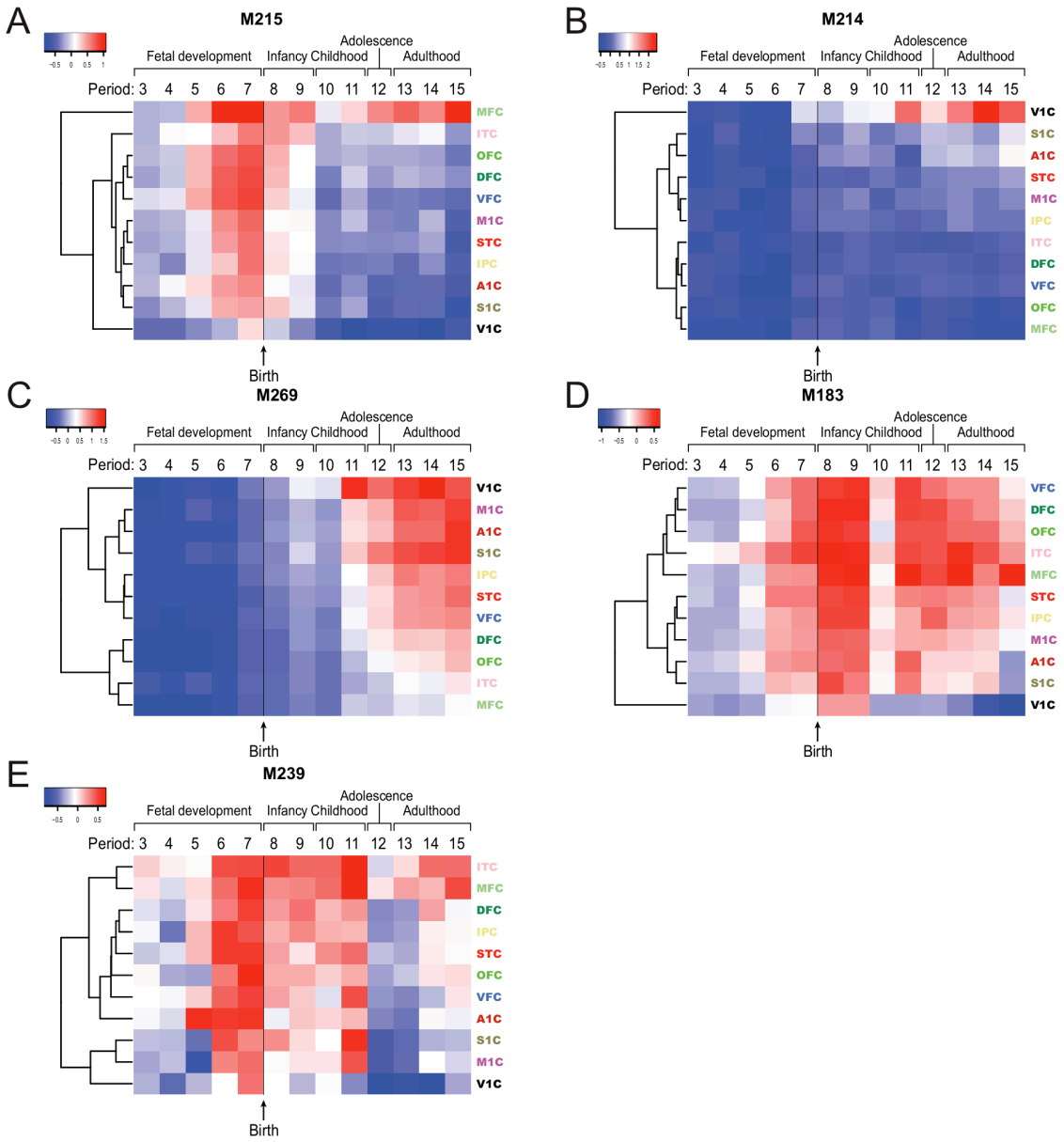


Figure S15 | Select adolescence and adulthood periods WGCNA modules.

Heat map of M215 (**A**) shows genes enriched postnatally in MFC (top genes: *GABRQ, SLN, CCBE1, SCN9A, C2orf85, ARHGAP6, CCDC109B, GRID2, SYT10,* and *HTR7*). M214 (**B**) genes are enriched postnatally in V1C (top genes: *TRPC3, ITGA11, KCNAB3, VAV3, SYT6, DPP4, EYA4, PTGS1, EEPD1,* and *SH3RF2*). M269 (**C**) genes are enriched in primary areas (top genes*: SYT2, SCN1B, SCN1A, HAPLN4, VAMP1, ANKRD29, SCN4B, RAB37, KCNS1,* and *STAC2*). M183 (**D**) genes are enriched in associative areas (top genes: *ANO3, PRKCG, SLIT1, SCN3B, RAB27B, FAM65B, SYT17, PRKCD, PCDH20,* and *PPP4R4*). M239 (**E**) genes are enriched in MFC and ITC (top genes: *ADAMTS19, CCDC148, DPYD, GHR, ARMC2, C2orf60, ANKRD7, KLHL1, COCH,* and *PPM1M*).

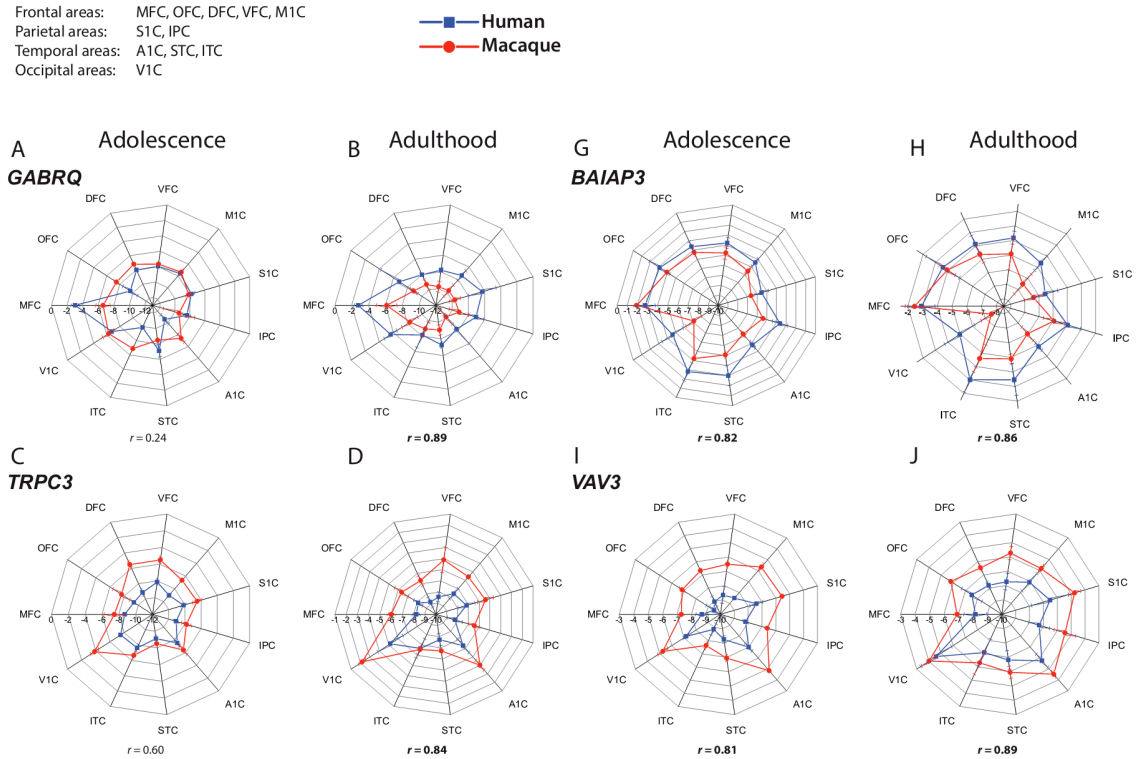


Figure S16 | Select adolescent and adult qRT-PCR validations.

Radar charts show the expression levels in all areas in adolescent and adult human and monkey brains. *GABRQ* exhibits enrichment in MFC in both adolescence and adulthood in human, but only in adulthood of the monkey (**A, B**). *TRPC3* shows enrichment in V1C in both adolescence and adulthood (**C, D**). *BAIAP3* shows enrichment in associative areas (**E, F**), and *VAV3* shows enrichment in primary sensory areas (**I, J**). Pearson's correlation coefficient *r* is printed bold when there is positive correlation between human and monkey pattern of expression. Human expression pattern is shown with blue lines and points while red is used for macaque.

2.9. References

1. K. Brodmann, *Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues* (Barth, Leipzig, 1909).
2. P. Rakic, *Science* **241**, 170 (1988).
3. M. Sur, J. L. Rubenstein, *Science* **310**, 805 (2005).
4. B. G. Rash, E. A. Grove, *Curr. Opin. Neurobiol* **16**, 25 (2006).
5. D. D. O'Leary, S. Sahara, *Curr Opin. Neurobiol* **18**, 90 (2008).
6. J. Hill *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13135 (2010).
7. C.-H. Chen *et al.*, *Science* **335**, 1634 (2012).
8. M. S. Gazzaniga, *Human: the science behind what makes us unique*. (Ecco, New York, 2008).
9. T. M. Preuss, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10709 (2012).
10. J.H. Lui, D. V. Hansen, A. R. Kriegstein, *Cell* **146**, 18 (2011).
11. J. H. Kaas, *Prog. Brain Res*. **195**, 91 (2012).
12. Z. Molnar, G. Clowry, *Prog. Brain Res.* **195**, 45 (2012).
13. H. Kennedy, C. Dehay, *Prog. Brain Res.* **195**, 341 (2012).
14. P. Broca, *Bulletin de la Société d'Anthropologie* **2**, 235 (1861).
15. M. S. Gazzaniga, R. W. Sperry, J. E. Bogen, *Proc. Natl. Acad. Sci. U.S.A.* **48**, 1765 (1962).
16. N. Geschwind, W. Levitsky, *Science* **161**, 186 (1968).
17. B. D. Cohen, C. D. Noblin, A. J. Silverman, S. B. Penick, *Science* **162**, 475 (1968).
18. R. A. Galuske, W. Schlote, H. Bratzke, W. Singer, *Science* **289**, 1946 (2000).
19. A. W. Toga, P. M. Thompson, *Nat. Rev. Neurosci.* **4**, 37 (2003).
20. G. Kasprian *et al.*, *Cereb Cortex* **21**, 1076 (2011).
21. T. Sun *et al.*, *Science* **308**, 1794 (2005).
22. K. Amunts, A. Schleicher, A. Ditterich, K. Zilles, *J. Comp. Neurol.* **465**, 72 (2003).
23. T. L. Hayes, D. A. Lewis, *Arch. Neurol.* **50**, 501 (1993).
24. J. G. Chi, E. C. Dooling, F. H. Gilles, *Arch. Neurol.* **34**, 346 (1977).
25. N. Geschwind, A. M. Galaburda, *Arch. Neurol.* **42**, 634 (1985).
26. P. Flechsig of Leipsic, *Lancet* **158**, 1027 (1901).
27. P. R. Huttenlocher, A. S. Dabholkar, *J. Comp. Neurol.* **387**, 167 (1997).
28. J. N. Giedd *et al.*, *Nat. Neurosci.* **2**, 861 (1999).
29. E. R. Sowell *et al.*, *Nat. Neurosci.* **6**, 309 (2003).
30. S. R. Chandana *et al.*, *Int. J. Dev. Neurosci*. **23**, 171 (2005).
31. D. C. Taylor, *Lancet* **2**, 140 (1969).

32. R. W. Thatcher, R. A. Walker, S. Giudice, *Science* **236**, 1110 (1987).
33. T. J. Cullen *et al.*, *Brit. J. Psychiat*. **188**, 26 (2006).
34. I. Voineagu *et al.*, *Nature* **474**, 380 (2011).
35. X. H. Piao *et al.*, *Science* **303**, 2033 (2004).
36. J. L. Rapoport, N. Gogtay, *Neuropsychopharmacology* **33**, 181 (2007).
37. H. Braak, E. Braak, J. Bohl, *Eur. Neurol.* **33**, 403 (1993).
38. H. J. Kang *et al.*, *Nature* **478**, 483 (2011).
39. M. B. Johnson *et al.*, *Neuron* **62**, 494 (2009).
40. M. J. Hawrylycz *et al.*, *Nature* **489**, 391 (2012).
41. S. J. Harrison-Uy, S. J. Pleasure, *Cold Spring Harb. Perspect. Biol.* **4**, (2012).
42. M. Tripodi, A. Filosa, M. Armentano, M. Studer, *Development* **131**, 6119 (2004).
43. S. A. Bayer, J. Altman, *Prog. Neurobiol.* **29**, 57 (1987).
44. C. Colantuoni *et al.*, *Nature* **478**, 519 (2011).
45. M. Dax, *Gaz. Heb. Méd. Chi.*, (1865).
46. I. Kostovic, M. Judas, Z. Petanjek, G. Simic, *Int. J. Psychophysiol.* **19**, 85 (1995).
47. K. Y. Kwan *et al.*, *Cell* **149**, 899 (2012).
48. F. Sanides, *Ann. N.Y. Acad. Sci.* **167**, 404 (1969).
49. M. H. Johnson, *Nat. Rev. Neurosci.* **2**, 475 (2001).
50. A. T. Kalinka *et al.*, *Nature* **468**, 811 (2010).
51. Y. Benjamini, Y. Hochberg, *J Roy Stat Soc B Met* **57**, 289 (1995).
52. P. Langfelder, S. Horvath, *Bmc Bioinformatics* **9**, (Dec 29, 2008).
53. P. Langfelder, S. Horvath, *J Stat Softw* **46**, 1 (Mar, 2012).
54. B. Zhang, S. Horvath, *Stat Appl Genet Mo B* **4**, (2005).
55. D. W. Huang, B. T. Sherman, R. A. Lempicki, *Nat Protoc* **4**, 44 (2009).
56. D. W. Huang, B. T. Sherman, R. A. Lempicki, *Nucleic Acids Res* **37**, 1 (Jan, 2009).
57. B. Clancy *et al.*, *Neuroinformatics* **5**, 79 (2007).
58. K. S. Saleem, N. Logothetis, *A combined MRI and histology atlas of the rhesus monkey brain in stereotaxic coordinates*. (Academic, London ; Burlington, MA, 2007), pp. ix, 326 p.

3. Multi-regional analysis of gene expression evolution in the human brain

André M. M. Sousa[1,2*], Ying Zhu[1*], Kyle A. Meyer[1], Yuka Imamura Kawasawa[1], Mingfeng Li[1], Marta Melé[3], Tomas Marques-Bonet[3], Nenad Šestan[1]

[1]Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510, USA.
[2]Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical Sciences Institute, University of Porto, Porto, Portugal.
[3]Institut de Biologia Evolutiva, Universitat Pompeu Fabra-Consejo Superior de Investigaciones Científicas, 08003 Barcelona, Spain

* These authors contributed equally to this work.
Correspondence:
Nenad Sestan, MD, PhD
E-mail: nenad.sestan@yale.edu

3.1. Abstract

Different brain regions possess distinct transcriptional profiles, and it has long been hypothesized that changes in transcriptome profiles have contributed to the evolution of the distinct phenotypical differences, including differences in cognitive and motor abilities, that exist between humans and non-human primates (NHPs). Here we report the RNA sequencing and analysis of 16 adult brain regions, including 11 neocortical areas, from six humans, five chimpanzees and five rhesus macaques. We found 6,389 genes were differentially expressed (DEX) among species, with 3,154 specifically up- or downregulated in humans. Furthermore, a multi-regional approach allowed us to examine intra-species DEX genes among all regions and compare how the transcriptional relatedness of areas differs among species. In addition, we found region-specific gene co-expression modules that are conserved across all three species, as well as species-specific transcription network modules, including one module that specifically discriminates human cortical areas. Several genes involved in the dopamine synthesis pathway were upregulated in humans, especially in the striatum. Finally, we analyzed the co-expression patterns of neurotransmitter receptors systems. Glutamatergic and GABAergic systems receptors were the most conserved among species, while serotoninergic and cholinergic systems were the least conserved. This study provides a comprehensive data set on the human and NHPs brain transcriptomes and insights into inter-primate and human-specific gene expression patterns.

3.2. Introduction

Comparisons of human and NHP genomes have yielded lists of coding (Enard, Khaitovich et al. 2002, Stedman, Kozyak et al. 2004, Wang and Su 2004, Vargha-Khadem, Gadian et al. 2005) and non-coding (Bejerano, Haussler et al. 2004, Pollard, Salama et al. 2006, Prabhakar, Noonan et al. 2006, Visel, Prabhakar et al. 2008) changes that may have been important in human brain evolution. Beyond genome level analyses, several studies have searched for expression differences between human and NHP brains (Somel, Franz et al. 2009, Babbitt, Fedrigo et al. 2010, Liu, Somel et al. 2012) an approach rooted in the idea that the evolution of regulatory elements, along with coding sequences, underlies many phenotypic differences between humans and NHPs (King and Wilson 1975). These previous studies provide useful information about expression differences in the human and NHP brain but also have several limitations. Most of the studies have used microarrays, reducing the comparison to identical probes among species and thus decreasing the number of genes compared. Furthermore, all studies have been restricted to small sample sizes comprising only a few brain regions due to difficulties in obtaining high-quality post-mortem human and NHP brain tissue. By expanding the focus to many regions, a more comprehensive understanding can be gained, as the brain consists of functionally divided and complexly interconnected regions and circuits. Studies on global transcription patterns in the human brain (Johnson, Kawasawa et al. 2009, Colantuoni, Lipska et al. 2011, Kang, Kawasawa et al. 2011) have revealed dynamic spatial regulation of transcription across different brain regions, including neocortical areas, further emphasizing the importance of analysing a wide range of brain regions. To better understand the evolution of the human brain, it is important to extend these multi-regional transcriptome studies to distinguish between features that conserved among primates and those that are specific to humans.

3.3. Results

Study design

We performed a comparative analysis of gene expression in 16 brain regions (Table 1) sampled from six human (*Homo sapiens*; H), five chimpanzee (*Pan troglodytes*; P), and five rhesus macaque (*Macaca mulatta*; M) adult specimens. RNA sequencing was carried out on a total of 247 samples (Supplementary Table 1), resulting in 4.6 billion uniquely mapped reads that were used to assess gene expression. Consistent and stringent quality control measures were applied at the level of tissue acquisition and dissection, sample preparation, and data processing (see the Supplementary Information for a detailed description). Briefly, samples were filtered by RIN number, sequencing error, and the number of uniquely mapped reads. A hierarchical clustering of all samples separated

samples from two macaque brains, which were prepared by different RNA extraction methods, from all others. This batch effect was corrected using *ComBat* (Supplementary Information, and Supplementary Fig. 1).

The macaque, which last shared a common ancestor with chimpanzees and humans around 25 million years ago (Ma; Sequencing, Consortium et al. 2007) served as an outgroup, enabling the inference of evolutionary species-specific differences between human and chimpanzees, which diverged approximately 6 Ma (Mikkelsen, Hillier et al. 2005). To infer the genetic distances among individuals and verify that samples are not more closely related than expected (for reasons such as inbreeding while in captivity), RNA-Seq reads were used to call SNPs and build neighbor-joining trees (Supplementary Information and Supplementary Fig. 2). Clustering reliably captured evolutionary relationships. As expected, humans had the lowest variation and macaques had the greatest.

In order to examine the contributions of different factors to transcriptional differences, we applied principal component analysis (PCA). The species to which each sample belonged (Figure 1a), rather than the region (Figure 1b), accounted for the most variance, indicating that inter-species differences are more pronounced than intra-species inter-regional differences in gene expression. Furthermore, the PC1, which accounts for 19.3% of the variability, separated macaque from human and chimpanzee samples, while the PC2, which explained 11.5% of the variance, separated human from chimpanzee samples (Figure 1).

Interspecies differential expression

We first focused on genes that had expression differences among species for each of the 16 brain regions. To avoid biases due to the quality of gene annotation for the three species, we generated a common annotation set by overlapping reciprocal mapping of human annotation onto the chimpanzee and macaque genomes (Supplementary Information) and considered only 1:1:1 species orthologs for all downstream analyses. For each region, DEX genes were calculated by comparing generalized linear models (GLM), with species as a main factor and batch as a co-factor.

We found 6,389 genes that were differentially expressed between at least two species in one or more regions. Based on *post hoc* comparisons (Supplementary Table 2), DEX genes were classified by the relationship among species, such as human and chimpanzee expression levels similar to each other and greater than macaque levels: H = P > M (Figure 2a). As expected, based on evolutionary relationships, the two most frequent patterns of differential expression had similar levels in human and chimpanzee

samples, which together differed from the expression levels in macaque samples (H = P > M and H = P < M).

Within all the different species relationships, regions had substantial differences in the number of DEX genes, highlighting the importance of a multi-regional analysis. For genes with human-specific up- or downregulation (H > P = M or H < P = M), STR showed an excess of DEX genes, indicating that this region may have played an important role in human brain evolution (Figure 2b). There has been conflicting evidence of a human-specific increase in gene upregulation (Caceres, Lachuer et al. 2003, Gu and Gu 2003, Khaitovich, Muetzel et al. 2004, Uddin, Wildman et al. 2004). In our study, 15 of the 16 regions showed a greater number of genes with human-specific upregulation than genes with human-specific downregulation, with STR being the only region with more human-specific downregulation. In contrast, chimpanzee brain regions were as likely to have a higher number of downregulated genes as upregulated, while only 2 macaque regions – HIP and STR – had more upregulated genes, with 14 showing more downregulated genes.

Among the genes with human-specific up- or downregulation (Supplementary Table 3), we found 31 genes that were DEX in all regions. Only 3 genes were significantly DEX (H > P = M or H < P = M) exclusively in the eleven neocortical areas: *TWIST1* (Figure 2c)*, RP11-364P22.1, and CTB-78F1.1*. There were also several genes that show human-specific DEX in only one brain region. Among those, we found that *ZP2*, a gene that was previously reported to be highly expressed in the postnatal human cerebellar cortex (Kang, Kawasawa et al. 2011) is not expressed, at our detection level, in any area/region of the chimpanzee and macaque brains (Figure 2d).

To gain some knowledge on which functional groups are more represented in our DEX gene list we performed gene ontology (GO) enrichment analysis. GO analysis on the species-specific DEX genes did not reveal any clear enrichment of functional categories, as most of the genes did not have distinctive functional annotation.

Intraspecies differential expression

In addition to species differences at the level of individual regions, we investigated how the regional transcriptional architecture differs among species. For each species, we calculated the number of intra-species DEX genes for all regions (Figure 3a) and for NCX areas only (Figure 3b). We found a smaller number of DEX genes in human brains regions, with macaque showing the highest number (Figure 3a). This pattern was especially pronounced in NXC areas (Figure 3b), indicating that the human NCX areas are very similar in terms of gene expression, when compared to chimpanzee or macaque. We also calculated pairwise correlation matrices of DEX genes and performed

hierarchical clustering of all regions to reveal the transcriptional relatedness of brain regions (Figure 3c-e). As expected (Lockhart and Barlow 2001, Roth, Hevezi et al. 2006), the CBC was the most divergent area of the brain in all species, followed by the STR and the MD (though in macaque the MD is more divergent from the cortical areas than the STR), thus clearly separating cortical areas from non-cortical ones. As mentioned above, NCX areas were more transcriptionally similar in humans than NCX areas were in either chimpanzees or macaques (Figure 3c-e). Among these areas, several intriguing species differences in clustering were present. Interestingly, the clearest difference in clustering was found in macaques, where all primary areas (M1C, S1C, A1C, and V1C) clustered together (figure 3e). Another noteworthy result was found in humans, which showed a clustering of IPC, STC, and VFC, a group of areas that comprise the perisylvian language areas of Broca and Wernicke (Broca 1861, Dax 1865, Wernicke 1874). This clustering was not present in the other two species.

Weighted gene co-expression network analysis (WGCNA)

As the DEX analysis focused on single genes and on a limited number of expression patterns, we next applied WGCNA to identify modules of highly co-expressed genes by searching for genes with similar patterns of variation across samples (Zhang and Horvath 2005). We identified 229 modules of co-expressed genes (Supplementary Table 4), many with distinct regional and/or species expression patterns (Figure 4a). Modules with patterns that are shared across all three species include M33, a module that is enriched in NCX areas (Supplementary Figure 3). Genes in this module are enriched in the categories “Homeobox, conserved site” (Bonferroni-adjusted P = 0.02) and “transcription factor" (Bonferroni-adjusted P = 0.02). Hub genes in this module include a handful of genes that have been linked to neocortical development, including *TBR1* (Hevner, Shi et al. 2001, Han, Kwan et al. 2011, McKenna, Betancourt et al. 2011), and *SATB2* (Britanova, de Juan Romero et al. 2008), two genes that have previously been shown to be co-regulated (Srinivasan, Leone et al. 2012). Species also shared several spatially restricted modules, such as M16 enriched in CBC (Supplementary Figure 4). Genes in M16 are associated with “transcription factor activity” (Bonferroni-adjusted P = 2.9E-4), “DNA-binding” (Bonferroni-adjusted P = 3.9E-4) and “developmental protein” (Bonferroni-adjusted P = 8.5E-4).

In contrast to these shared modules, other modules were distinguished by species differences. Among human-specific modules, M81, 162, 192, 229, were upregulated in all human regions (Supplementary Figures 5-8). Similarly, M215 is selectively increased in human cortical regions (Figure 4b). Genes in M81 were associated with “TSPN” (Bonferroni-adjusted P = 0.02), in M192 with “oxidative phosphorylation” (Bonferroni-

adjusted *P* = 0.03), and in M229 with "ribosome" (Bonferroni-adjusted *P* = 1.2E-3), "ribosomal subunit" (Bonferroni-adjusted *P* = 0.01), and "large ribosomal subunit" (Bonferroni-adjusted *P* = 0.02). No significant pathways were identified for M162 or M215 after Bonferroni adjustment. There were modules that show human-specific decreases in all regions, such as M69, and M173 (Supplementary Figures 9 and 10). No significant pathways were identified for M69 and M173 after Bonferroni adjustment.

Interestingly, some modules revealed species differences in only a small subset of regions. M130 is composed of genes that were upregulated in human AMY, HIP, and STR, which is the region with the most interspecies DEX genes (Figure 4d). Intriguingly, 2 hub genes in this module are involved in the synthesis of dopamine: *TH* and *DDC* in M130 (Figure 4e). One of the hub-genes in the M215 (Figure 4c), *PAH*, may also be part of the dopamine biosynthesis pathway. Due to this finding, we carried out a more detailed analysis of genes involved in dopamine metabolism.

Human-specific expression of genes involved in dopamine synthesis

Dopamine is a monoamine neurotransmitter that belongs to the catecholamine family, and it is the precursor of epinephrine and norepinephrine. Dopamine is synthetized from either L-phenylalanine or L-tyrosine by three enzymes in consecutive steps: phenylalanine hydroxylase (PAH) converts L-phenylalanine into L-tyrosine; tyrosine hydroxylase (TH) converts L-tyrosine into L-dihydroxyphenylalanine (L-DOPA); finally, DOPA decarboxylase (DDC) converts L-DOPA into dopamine (Figure 5a). We revealed that *PAH* exhibited higher expression levels in the human neocortical areas, when compared to chimpanzee and macaque (Figure 5b). More importantly, we found that *TH* and *DDC* showed higher expression levels in the STR of humans (Figures 5d and 5f). Furthermore, the genes coding for two of the enzymes involved in dopamine degradation, *COMT* and *MAOA*, did not show any significant differences, in expression, between species (Supplementary Figure 11). All these findings were validated using droplet digital PCR (ddPCR; Figure 5c, 5e, and 5g). Together, this indicates that humans may have higher levels of dopamine in STR, because humans are producing more dopamine, which is then presumably degraded at the same level as in the other species. To determine if these differences were reflected at the protein level, we performed protein immunoblotting of STR lysates from human, chimpanzee, and macaque samples. Remarkably, both TH and DDC showed higher levels in human STR than in the other species (Supplementary Figure 12). The difference was more pronounced in TH, which is the enzyme that catalyzes the rate-limiting step in the biosynthesis of dopamine.

We also analyzed dopamine receptors expression patterns to verify if the upregulation of dopamine production is accompanied by an upregulation of dopamine

receptors. Surprisingly, we found 3 genes that encode dopamine receptors – *DRD1*, *DRD2*, and *DRD3* – were downregulated in the human STR (Supplementary Figure 13), which may have a balancing effect on the overproduction of dopamine in the human STR.

These differences in dopamine biosynthesis and neuroreceptors encouraged us to extend our analyses to other neurotransmitter systems.

Neurotransmitter receptors networks

Although differences in neurotransmitter biosynthesis may alter physiological properties of neural circuits because the levels of neurotransmitters will affect synaptic transmission, the composition of receptors is also relevant. We generated co-expression networks between neuroreceptor subunits and measured the conservation of these networks among the three species. A statistic Z was calculated to evaluate the preservation of the network compared with randomly selected sets containing the same number of genes (Langfelder et al, 2011). We limited the analysis to four systems that involved enough genes to provide reliable comparisons: glutamatergic, GABAergic, serotoninergic, and cholinergic (Figure 6). The major excitatory and inhibitory systems in the brain, i.e. the glutamatergic and the GABAergic systems (Figure 6a and 6b, respectively), have similar correlation networks among species (Supplementary Figure 14). This conservation is unsurprising given that these two systems are ubiquitous and have a vital role in brain function. On the other hand, the two neuromodulatory systems analyzed, i.e. the cholinergic and the serotoninergic systems, have significantly different networks: human and chimpanzee serotoninergic networks are different from macaque (Figure 6c and Supplementary Figure 14); and chimpanzee have a significantly different cholinergic network when compared to human and macaque (Figure 6d and Supplementary Figure 14).

3.4. Conclusion

Our study uncovers several human-specific features at the level of gene expression, allowing a better understanding of some aspects of human brain evolution. We show that some areas of the brain, like the striatum, show more differences between species than others. Species also exhibit different transcriptional architectures between brain regions. Additionally, all species show specific network modules of co-expressed genes, including two specific modules with human-specific expression of genes in cortical areas. We also found remarkable differences in the dopamine synthesis pathway, with several important genes being upregulated in humans. Moreover, we show differences among neurotransmitter receptors and reveal that some of them have a more conserved expression pattern than others. This data set (available at

www.humanbraintranscriptome.com) provides a basis for further functional studies on genes that might be of high relevance for human-specific traits and also for future comparisons with other transcriptome data sets.

Our study expands the knowledge of spatial differences in gene expression that distinguish the human brain from other NHPs, but uncovering how these regional differences vary throughout time, especially during prenatal and postnatal development, will be essential for a better understanding of human brain development and evolution.

3.5. Methods summary

Supplementary Information provides a full description of tissue acquisition and processing, data generation, validation and analyses.

3.6. Acknowledgements We thank M. Horne, Chet Sherwood, and Patrick Hof for assistance with tissue acquisition, and Darshani Singh and Yurae Shin for technical assistance. We also thank members of the Sestan laboratory for discussions and criticism. Support for predoctoral fellowships was provided by the Portuguese Foundation for Science and Technology (A.M.M.S.), the China Scholarship Council (Y.Z.), and National Science Foundation Graduate Research Fellowship (K.A.M). This work was supported by grants from the US National Institutes of Health (MH081896, MH089929, NS054273), the Kavli Foundation and NARSAD, and by a James S. McDonnell Foundation Scholar Award (N.S.).

3.7. Tables

Table 1 | Specimens list

Number	Species	Sex	Age (years)	Hemisphere
HSB 123	*Homo sapiens*	Male	37	Right
HSB 126	*Homo sapiens*	Female	30	Right
HSB 130	*Homo sapiens*	Female	21	Left
HSB 135	*Homo sapiens*	Female	40	Right
HSB 136	*Homo sapiens*	Male	23	Right
HSB 145	*Homo sapiens*	Male	36	Right
PTB 162	*Pan troglodytes*	Female	23	Left
PTB 164	*Pan troglodytes*	Female	31	Right
PTB 165	*Pan troglodytes*	Male	31	Right
PTB 166	*Pan troglodytes*	Male	27	Right
PTB 167	*Pan troglodytes*	Male	30	Right
RMB 160	*Macaca mulatta*	Female	11	Left
RMB 161	*Macaca mulatta*	Male	11	Left
RMB 196	*Macaca mulatta*	Female	11	Right
RMB 218	*Macaca mulatta*	Male	7	Left
RMB 219	*Macaca mulatta*	Male	7	Left

3.8. Figures

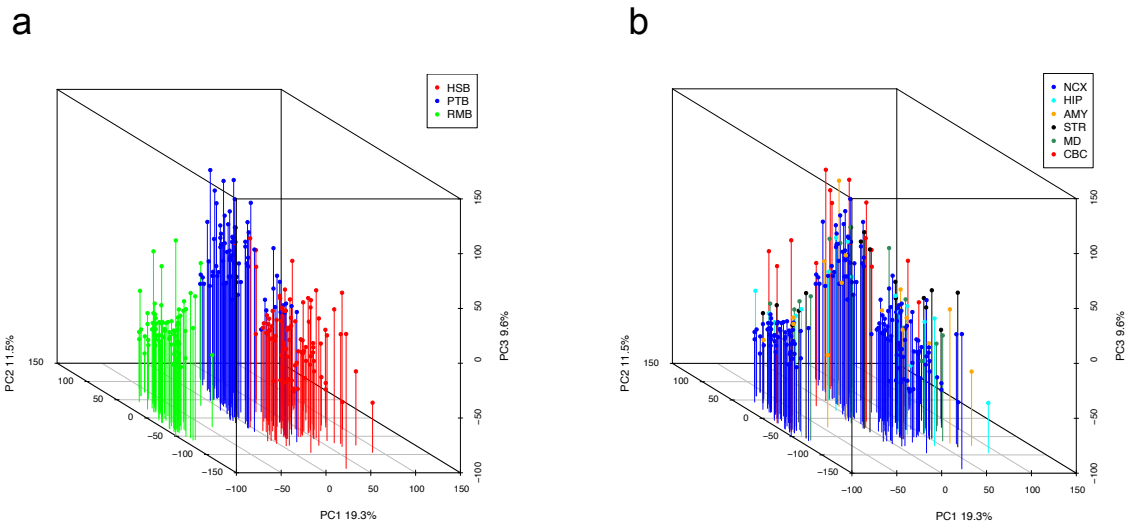


Figure 1. Principal component analysis

Three-dimensional plot of PCA of all samples. Each point represents one sample. Samples are colored by species (**a**) or by region (**b**). The percentage of variation that is explained by each Principal Component is shown.

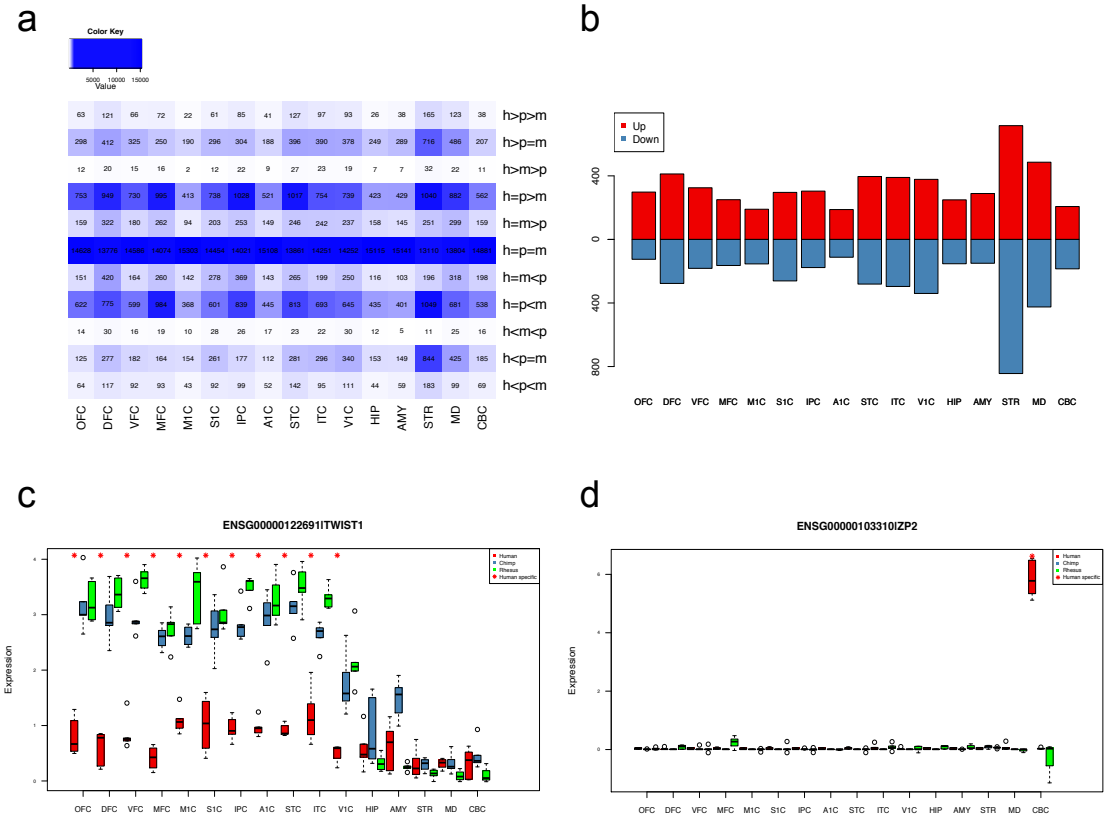


Figure 2. Interspecies differential expression

a, Heat map showing the number of genes differentially expressed, according to *post hoc* comparisons explained in supplementary table 2. **b**, Number of human-specific genes that are up- (h>p=m; red) or downregulated (h<p=m; blue) across the 16 brain regions. **c**, Expression [$\log_2$ (RPKM+1)] pattern of *TWIST1*. Significant differences are labeled with an asterisk. **d,** Expression pattern of *ZP2.* Significant differences are labeled with an asterisk.

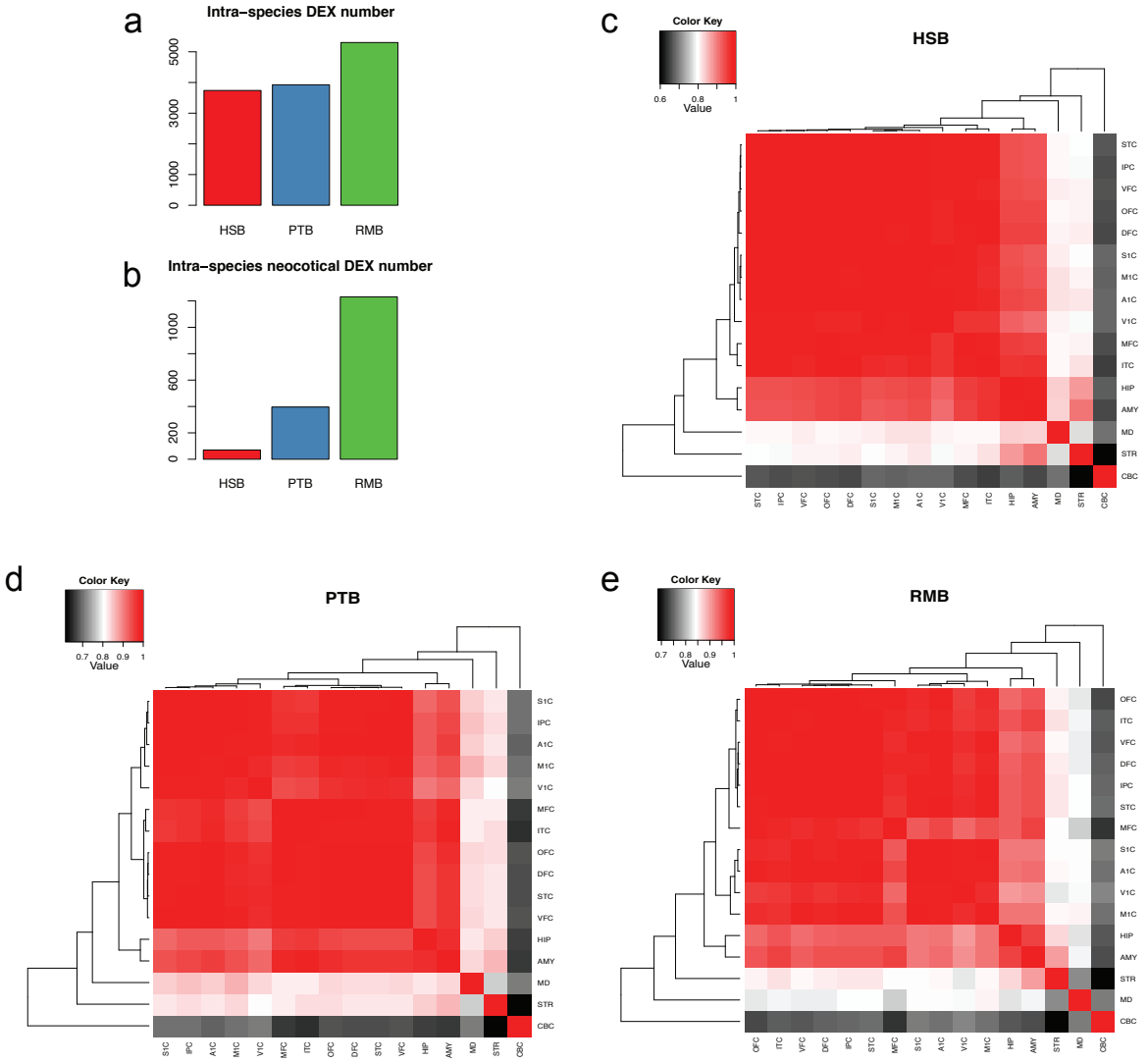


Figure 3. Intraspecies differential expression

a, Bar plot showing the number of intra-species DEX genes across all regions in human (red), chimpanzee (blue), and macaque (green). **b**, Bar plot showing the number of intra-species DEX genes across the neocortical areas in human (red), chimpanzee (blue), and macaque (green). **c, d, e,** Heat maps of the unsupervised clustering of all regions in human, chimpanzee, and macaque, respectively.

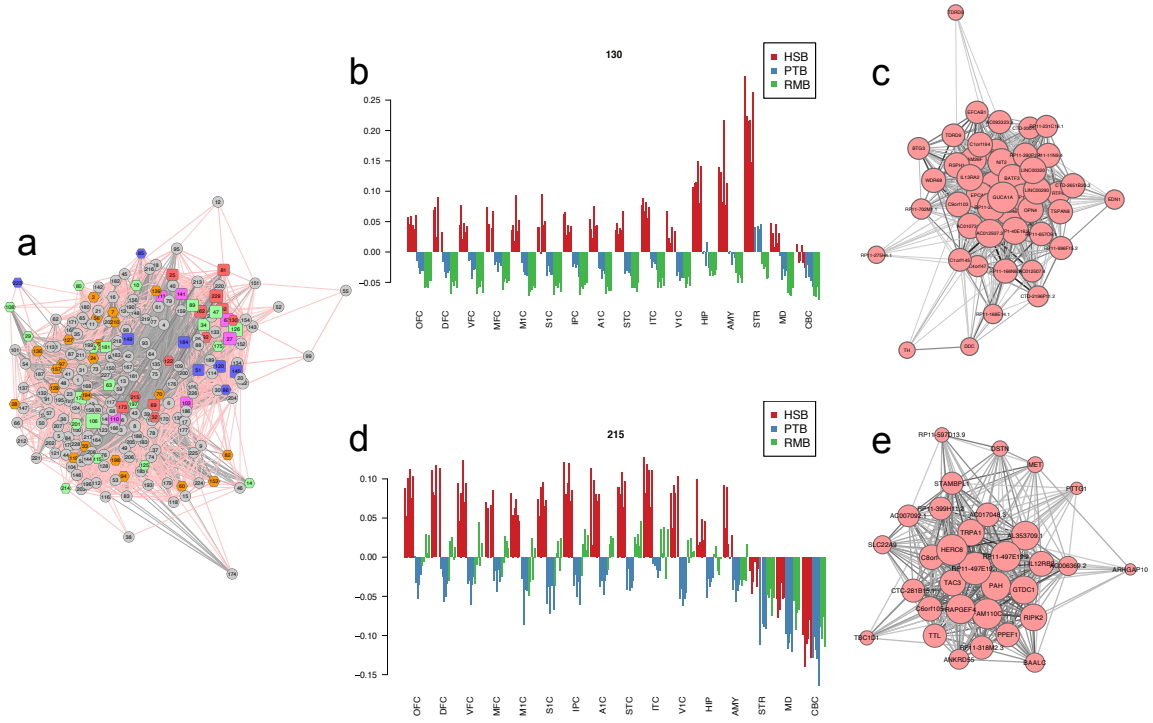


Figure 4. Global co-expression networks and gene modules

a, Figure showing the gene co-expression network analysis of all samples. Modules labeled as human-specific (red), chimpanzee-specific (blue), macaque-specific (green). Modules with a conserved expression (orange) and in a gradient pattern (violet) among species are labeled. Modules represented by a square include genes that are co-expressed in all regions. Modules with a hexagon include genes that are co-expressed in only one region. **b**, Bar plot showing the spatial expression pattern of module 130. **c**, M130 genes. **d**, Bar plot showing the spatial expression pattern of module 215. **e**, Module 215 genes.

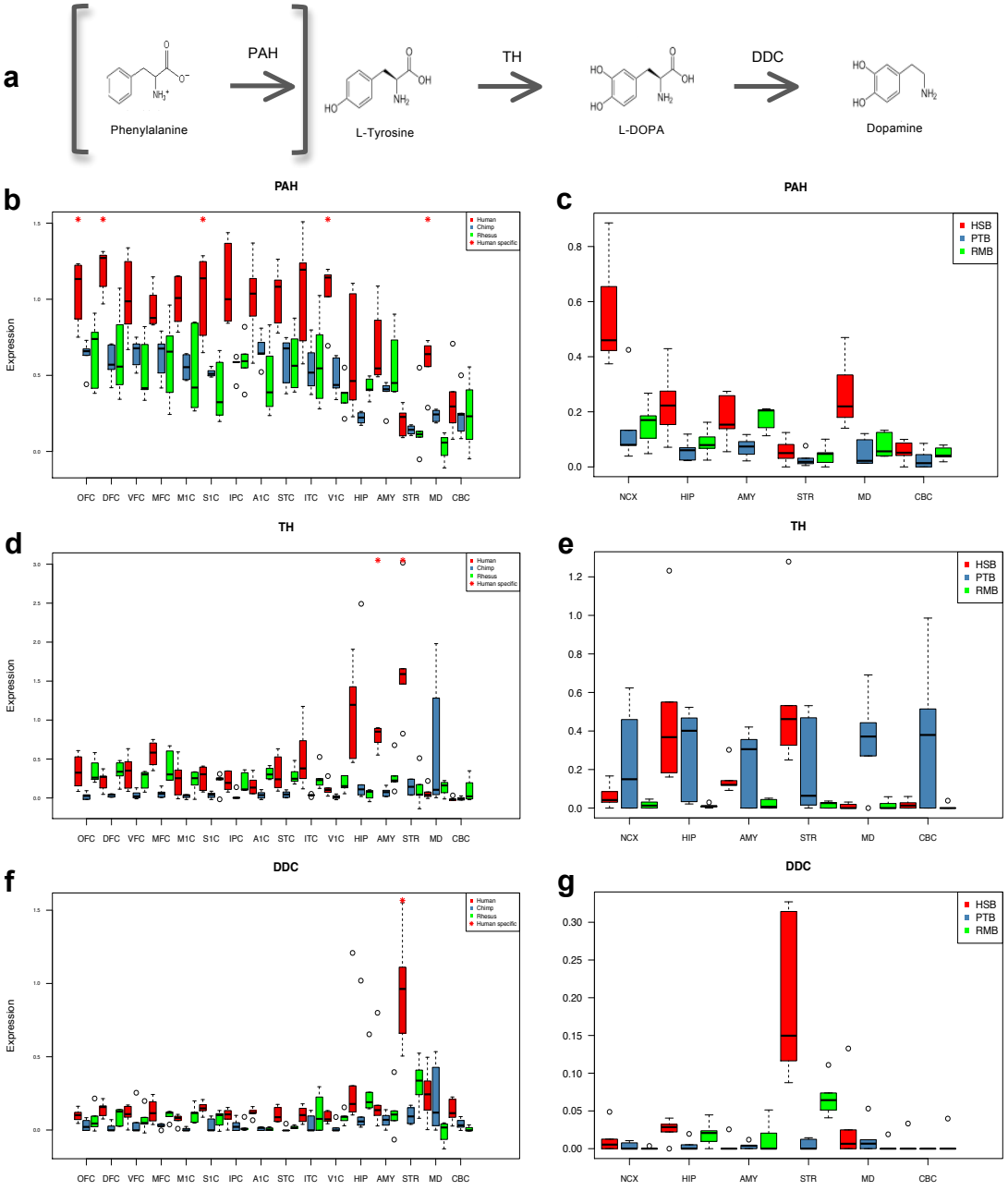


Figure 5. Dopamine biosynthesis pathway

a, Schematic figure of the dopamine biosynthesis pathway. **b**, Expression [$\log_2$ (RPKM+1)] pattern of *PAH*. This gene is significantly upregulated in several human neocortical áreas and the MD. **c**, ddPCR validation of *PAH* expression normalized to the housekeeping gene TBP. **d**, Expression pattern of *TH*. This gene is significantly upregulated in human striatum and amygdala. **e**, ddPCR validation of *TH* expression normalized to the housekeeping gene *TBP*. **f**, Expression pattern of *DDC*. This gene is significantly upregulated in the human striatum. **g**, ddPCR validation of *DDC* expression normalized to the housekeeping gene *TBP*.

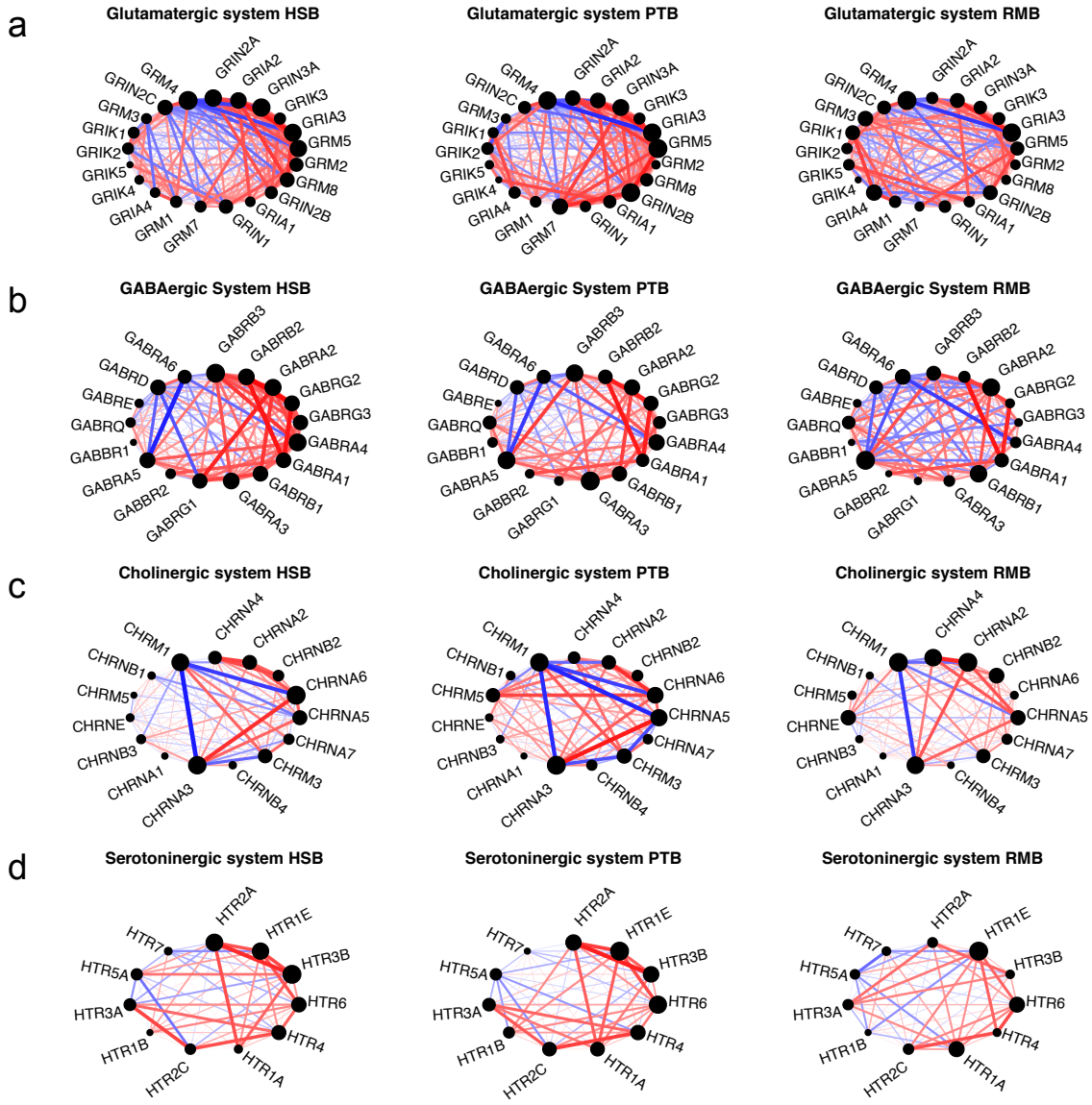


Figure 6. Neurotransmitter receptor systems

Network of co-expression of genes encoding glutamatergic system (**a**), GABAergic (**b**), cholinergic (**c**), and serotoninergic (**d**) receptor subunits. Positive correlation (red) and negative correlation (blue) between genes are shown.

3.9. Supplementary Information

Table of Contents

1. Introduction

2. Ontology and anatomical definition of sampled brain regions and NCX areas

3. Tissue Procurement and Sampling

3.1 Tissue procurement

3.2 Neuropathological evaluation

3.3 Selection criteria for brain specimens

3.4 Tissue dissection

3.4.1 Regional sampling from fresh brain specimens

3.4.2 Regional sampling from frozen brain specimens

3.5 Histological verification of tissue sampling

3.6 Dissection scoring

3.7 Tissue pulverization

4. RNA Isolation, Library Preparation, Sequencing, and Quality Assessment

4.1 RNA extraction

4.2 Library preparation for mRNA-sequencing

4.3 Sequencing

5. mRNA Sequencing Data Analysis

5.1 mRNA sequencing alignment, annotation, and expression quantification

5.1.1 Reference genomes and read alignment

5.1.2 Generation of inter-species annotation

5.1.3 Expression quantification

5.2 Quality control assessment and detection of outliers

5.2.1 Number of uniquely mapped reads

5.2.2 Uniformity of reads' coverage along the annotated genes

5.2.3 Hierarchical clustering and batch effect correction

5.2.4 Principal component analysis

5.2.5 Effects of RIN and PMI on mRNA expression

5.3 Identification of DEX genes

5.3.1 Identification of intra-species DEX genes

5.3.2 Identification of inter-species DEX genes

5.4 Weighted gene co-expression network analysis (WGCNA)
5.5 Conservation of neurotransmission systems

6. Functional enrichment analysis

7. Determination of genetic relationships among samples
7.1 Mapping
7.2 Identification of single nucleotide variants
7.3 Construction of phylogenetic tree

8. Validation of dopamine metabolism genes
8.1 Droplet digital PCR
8.2 Immunoblotting

9. Supplementary Tables
Supplementary Table 1 | Tissue sample metadata
Supplementary Table 2 | Models used to categorize interspecies DEX genes
Supplementary Table 3 | Human-specific DEX genes
Supplementary Table 4 | Weighted co-expression network modules
Supplementary Table 5 | Spike-in RNAs
Supplementary Table 6 | List of PCR primers used for ddPCR validation

10. Supplementary Figures
Supplementary Fig. 1 | Unsupervised sample clustering
Supplementary Fig. 2 | Species neighbor-joining tree from variant calling
Supplementary Fig. 3 | WGCNA Module 33
Supplementary Fig. 4 | WGCNA Module 16
Supplementary Fig. 5 | WGCNA Module 81
Supplementary Fig. 6 | WGCNA Module 162
Supplementary Fig. 7 | WGCNA Module 192
Supplementary Fig. 8 | WGCNA Module 229
Supplementary Fig. 9 | WGCNA Module 69
Supplementary Fig. 10 | WGCNA Module 173
Supplementary Fig. 11 | Expression pattern of genes involved in dopamine degradation
Supplementary Fig. 12 | Immunoblots of TH and DDC in human

Supplementary Fig. 13 | Expression pattern of genes coding dopamine receptors
Supplementary Fig. 14 | Neurotransmitter networks conservation

1. Introduction

In this Supplementary Information we provide further information regarding the study design, materials and methods, and additional data. The materials and methods section provide detailed description of the collection, dissection methods, and quality control assessments of human, chimpanzee and macaque postmortem tissue used in this study. We provide technical descriptions of data generation and analyses using different platforms. We also make available additional data and supplementary figures and tables that were discussed in the main manuscript. Some information for the human samples and procedures is shared with our previous study (Kang, Kawasawa et al. 2011) and is reproduced for completeness.

2. Ontology and anatomical definition of sampled brain regions and NCX areas

In the present study, we analysed the same 16 brain regions as in our previous study (Kang, Kawasawa et al. 2011). All brains were dissected using the same protocol used for human specimens. Chimpanzee and macaque brain regions were sampled according to DeLucchi (de Lucchi, Dennis et al. 1965) and Saleem and Logothetis (Saleem and Logothetis 2007) reference atlases, respectively.

3. Tissue Procurement and Sampling

3.1 Tissue procurement

Postmortem human brain specimens were obtained from tissue collections at the Department of Neurobiology at Yale University School of Medicine and the Clinical Brain Disorders Branch of the National Institute of Mental Health [human paper ref?]. Tissue was collected after obtaining parental or next of kin consent and with approval by the institutional review boards at the Yale University School of Medicine, the National Institutes of Health, and at each institution from which tissue specimens were obtained. Tissue was handled in accordance with ethical guidelines and regulations for the research use of human brain tissue set forth by the NIH (http://bioethics.od.nih.gov/humantissue.html) and the WMA Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html). Appropriate informed consent was obtained and all available non-identifying information was recorded for each specimen. Specimens range in age from 21 to 40 years. The postmortem interval (PMI) was defined as hours between time of death and time when tissue samples were frozen.

All experiments using nonhuman primates were carried out in accordance with a protocol approved by Yale University's Committee on Animal Research and NIH guidelines.

3.2 Neuropathological evaluation

All clinical histories, tissue specimens, and histological sections were evaluated to assess for hypoxia, cerebrovascular incidents, tumours, microbial infections, neurodegeneration, demyelination, and metabolic disease. To prepare tissue sections for microscopic neuropathological examination, small samples (usually the dorsal parietal cortex, striatum with ependymal layer and subependymal zone, hippocampus, and the cerebellar cortex) were dissected and fixed in 4% paraformaldehyde and processed for histology and immunohistochemistry as described below.

3.3 Selection criteria for brain specimens

To better ensure consistency between samples and decrease potential variation due to ante- and postmortem conditions, specific selection criteria were arbitrarily established. The aim was to collect tissue specimens from clinically unremarkable donors. The following selection criteria were strictly adhered to when deciding whether to exclude or include each brain specimen:

- Only brains free of obvious malformations or lesions were collected. Disqualifying characteristics included prominent intraparenchymal haemorrhage and ischemia, infection, periventricular leukomalacia, abnormal meninges, dysplasia, hypoplasia, alterations in the pial or ventricular surface, and extensive white matter lesions.
- Samples were excluded if microscopic analysis revealed extensive neuronal loss, neuronal swelling, glioneuronal heterotopias, or dysmorphic neurons and neurites.
- Samples that tested positive for Hepatitis B, Hepatitis C, or HIV were excluded.
- Human specimens were excluded if excessive drug or alcohol abuse was reported. All specimens with any known neurological or psychiatric disorders, or any prolonged agonal conditions (coma, prolonged pyrexia, hypoxia, seizures, prolonged dehydration, hypoglycaemia, and multiple organ failure) were not included in this study. Other excluding factors included ingestion of neurotoxic substances at the time of death, suicide (within humans), severe head injury, significant haemorrhages, widespread vascular abnormalities, ischemia, tumours, stroke, congenital neural abnormalities, and signs of neurodegeneration (spongiosis, amyloid plaques, neurofibrillary tangles, Lewy bodies, and amyloid angiopathy).

3.4 Tissue dissection

Different dissection procedures were used for each specimen depending on whether the samples were frozen (human and chimpanzee specimens) or fresh (macaque specimens).

3.4.1 Regional sampling from fresh brain specimens

Brains were chilled on ice for 15–30 minutes prior to sectioning. Brains were placed ventral side up onto a chilled aluminium plate (1 cm thick) on ice. The brainstem and cerebellum were removed from the cerebrum by making a transverse cut at the junction between the diencephalon and midbrain. Next, they were sectioned to obtain 1 cm (period 5–15 specimens) or 0.5 cm (periods 3 and 4 specimens) thick serial, coronal sections. The tissue slabs were snap frozen in isopentane (J.T. Baker)/dry ice at -30 to -40 ºC and stored at -80 ºC.

3.4.2 Regional sampling from frozen brain specimens

All previously frozen periods 3–15 specimens and tissue slabs were microscopically inspected and the desired region was demarcated, then dissected using a dental drill (AnyXing, 300D) and a Lindemann Bone Cutter H162A.11.016 or diamond disk saw (Dental Burs USA; r=11 mm) on a 1 cm thick aluminium plate over dry ice. Dissected tissue samples were stored at -80 ºC prior to further processing.

3.5 Histological verification of tissue sampling

To verify that the region or NCX area of interest is properly and consistently sampled, we also collected small tissue blocks, adjacent to the tissue sample dissected for the RNA extraction. We have done this for the majority of M1C, S1C, IPC, A1C and V1C samples, which in our experience were hard to match across different specimens but can be histologically verified using Nissl method in postnatal specimens due to cytoarchitectonic differences. This method was also occasionally used for other regions or NCX areas. These tissue blocks were then fixed in 4% paraformaldehyde for 48 h, sectioned at 50 µm thickness using a vibratome, and Nissl stained to verify the identities of dissected adjacent tissue (data not shown).

3.6 Tissue pulverization

To ensure proper representation of the region of interest, frozen tissue samples were pulverized in liquid nitrogen using a ceramic mortar and pestle (Fisher Scientific, cat# 12-961C and 12-961-5C). Pulverized samples were transferred to chilled wide-mouth

cryogenic vials (Nalgene, cat# 03-337-7B) and stored at -80 ºC until used for RNA extraction.

4. RNA Isolation, Library Preparation, Sequencing, and Quality Assessment

4.1 RNA extraction

A bead mill homogenizer (Bullet Blender, Next Advance) was used to lyse the pulverized tissue. Each pulverized tissue sample was transferred to a chilled safe-lock microcentrifuge tube (Eppendorf). A mass of chilled stainless steel beads (Next Advance, cat# SSB14B) equal to the mass of the tissue was added to the tube. Two volumes of lysis buffer were added to the tissue and beads. Samples were mixed in the Bullet Blender for 1 min at a speed of six. Samples were visually inspected to confirm desired homogenization and then incubated at 37 °C for 5 min. The lysis buffer was added up to 0.6 ml, and samples were mixed in the Bullet Blender for 1 min. Total RNA was extracted using RNeasy Plus Mini Kit (Qiagen) for mRNA-sequencing and using mirVana kit (Ambion) for small RNA-sequencing. Total RNAs from two macaques (RMB 218 and RMB 219) were also extracted using mirVana kit. Each sample was subjected to a DNase treatment (TURBO DNase, Ambion) as per manufacturer's instructions.

Optical density values of extracted RNA were measured using NanoDrop (Thermo Scientific) to confirm an A260:A280 ratio above 1.9. RIN was determined for each sample using Bioanalyzer RNA 6000 Nano Kit (Agilent), depending upon the total amount of RNA.

4.2 Library preparation for mRNA-sequencing

cDNA libraries were prepared using the mRNA-Seq Sample Kit (Illumina) as per the manufacturer's instructions with some modifications. Briefly, polyA RNA was purified from 1 to 5 µg of total RNA using (dT) beads. Quaint-IT RiboGreen RNA Assay Kit (Invitrogen) was used to quantitate purified mRNA with the NanoDrop 3300. Following mRNA quantitation, 2.5 µl spike-in master mixes, containing five different types of RNA molecules at varying amount (2.5×10^{-7} to 2.5×10^{-14} mol), were added per 100 ng of mRNA (Jiang et al., 2011). The spike-in RNAs were synthesized by External RNA Control Consortium (ERCC) consortium by in vitro transcription of de novo DNA sequences or of DNA derived from the *B. subtilis* or the deep-sea vent microbe *M. jannaschii* genomes and were a generous gift of Dr. Mark Salit at The National Institute of Standards and Technology (NIST). These were used both to track the brain regions and to normalize expression levels across experiments. Each sample was tagged by adding a pair of spike-in RNAs unique to the region from which the sample was taken. Also, an additional three common spike-ins were added for controlling sequencing error rates, which is not influenced by SNP existence (Supplementary Table 5). The mixture of mRNA and spike-in

RNAs were subjected to fragmentation, reverse transcription, end repair, 3'– end adenylation, and adapter ligation to generate libraries of short cDNA molecules. The libraries were size selected at 200 – 250 bp by gel excision, followed by PCR amplification and column purification. For RMB 218 and 219, we used TruSeq RNA Sample Prep Kit to generate cDNA libraries as per the manufacturer's instructions with the similar modification of adding spike-in RNAs before the mRNA fragmentation. The TruSeq Kit allows addition of 12 different types of index, thus the multiplexing of samples when they are sequenced. The final product was assessed for its size distribution and concentration using Bioanalyzer DNA 1000 Kit.

4.3 Library preparation for small RNA-sequencing

The TruSeq Small RNA Sample Kit (Illumina) was used to prepare cDNA libraries per manufacturer instructions. Briefly, 1 µg of total RNA was ligated with 3'- and then 5'-adapters, followed by reverse transcription and PCR amplification. The PCR utilizes 48 different types of primer that will add 48 different index sequences to the adapters. Each library was assessed for the presence of desired micro RNA population by Bioanalyzer High Sensitivity DNA Kit. We pooled 21 samples (16 samples from one brain and 5 technical replicates) with distinct indexes, which allow subsequent retrieval of each sample from multiplexed sequencing runs. Each pooled library was size selected by gel excision. The final product was assessed for its size distribution and concentration using Bioanalyzer DNA 1000 Kit.

4.4 Sequencing

We used Illumina Genome Analyzer IIx (GAIIx) for mRNA-sequencing by loading one sample per lane. But for RMB 218 and 219, we used HiSeq 2000 by loading 4 samples per lane. For small RNA-sequencing, we used Illumina HiSeq 2000 by loading 21 samples per lane.

For GAIIx sequencing, the library was diluted to 10 nM in EB buffer and then denatured using the Illumina protocol. The denatured libraries were diluted to 12 pM, followed by cluster generation on a single-end Genome Analyzer IIx (GAIIx) flow cell (v4) using an Illumina cBOT, according to the manufacturer's instructions. The Illumina GAIIx flow cell was run for 75 cycles using a single-read recipe (v4 sequencing kits) according to the manufacturer's instructions.

For HiSeq 2000 sequencing, the library was diluted to 10 nM in EB buffer and then denatured using the Illumina protocol. The denatured libraries were diluted to 15 pM, followed by cluster generation on a single-end HiSeq flow cell (v1.5) using an Illumina cBOT, according to the manufacturer's instructions. The HiSeq flow cell was run for 75

cycles for RMB218 and 219 mRNA-sequencing, and for 50 cycles for small RNA-sequencing using a single-read recipe (v2 sequencing kit) according to the manufacturer's instructions.

5. mRNA Sequencing Data Analysis

5.1 mRNA sequencing alignment, annotation, and expression quantification

5.1.1 Reference genomes and read alignment

The exporter files generated from Illumina's CASAVA pipeline were converted to fastq file. Reads were then filtered by quality score and mapped to the human (*Homo sapiens*, Feb. 2009 hg19), chimpanzee (*Pan troglodytes*, Oct. 2010 panTro3), and rhesus macaque genomes (*Macaca mulatta*, Jan. 2006 rheMac2), as appropriate, using TopHat without providing junction annotation.

5.1.2 Generation of inter-species annotation

To calculate the gene expression, we first created a composite gene model, using RSEQtools (Habegger et al., 2010), based on human Gencode v10 (Harrow et al., 2012). In the composite model, the overlapped exons in the same gene were merged to one to avoid double counting.

To find the non-human primate orthologous sequences for human exons, we mapped the human composite exons using liftOver (Hinrichs et al., 2006) followed by multiple filtering steps using BLAT (Kent 2002).

LiftOver: The coordinates of exons on the human genome were converted to those on the chimpanzee and the rhesus macaque genomes, according to the instructions provided in liftOver. The parameter –minMatch was set to 0.98 for chimpanzee and 0.913 for macaque, respectively, based on exon sequence similarity simulation (data not shown). The resulting coordinates were then converted back to the human genome, with the same –minMatch setting. The exons that did not map to the same human genomic location and the genes with exons mapping to different chromosomes or different strands were excluded from further analysis.

BLAT: As liftOver is based on alignments of big stretches, we performed multiple steps of filtering with BLAT to obtain an accurate conversion.

- Exons with low identity score (ID) and low percentage of mapped length (PL) were excluded. IDs were set to 0.95 and 0.9 for human-chimp and human-macaque exon conversion, respectively. PL was set to 90%.

- To avoid the difficulties in estimating multi-mapped reads, we further filtered the exons (ID > 0.95 for human-chimp and ID > 0.9 for human-macaque, PL > 90%) and genes (ID > 95%, PL > 80%) with highly similar duplicated sequences.
- To ensure the remaining sequences represent the original gene model, we kept only genes with length > 500bps or with at least 1/3 of the original length. Besides, we compared the rpkm of human genes using the converted model with that using original Gencode v10 annotation and filtered out genes with mean absolute difference greater than 1.

5.1.3 Expression quantification

After the reads were mapped to the reference sequences, the expression level of genes and exons were quantified by count number and RPKM (reads per kilobase of exon model per million mapped reads) (Mortazavi, Williams et al. 2008), using RSEQtools. Only genes with count number greater than 10 and maximum regional mean RPKM greater than 1 were included in further analysis.

5.2 Quality control assessment and detection of outliers

Several quality control measures were implemented throughout sample preparation and transcriptome data generation steps as described in Kang et al. 2011. Samples that failed to pass the quality control measures were reprocessed or removed from the downstream analyses.

5.2.1 Number of uniquely mapped reads

The number and the percentage of uniquely mapped reads were calculated for each sample. Samples with less than 5 million reads or low percentage of uniquely mapped reads were excluded or re-sequenced.

5.2.2 Uniformity of reads' coverage along the annotated genes

All genes were evenly divided into 100 segments from 5' to 3' end. We calculated the number of nucleotides mapped to each segment. Only genes with more than 1000 nucleotides mapped were used. The percentage of nucleotides (PN) mapped to each segment was calculated as the number of nucleotides in each segment divided by the total number of nucleotides mapped to the whole gene. The median of segmental PN for all genes was calculated to represent segmental PN of a specific sample. The

quartiles of sample segmental PN were then plotted, with solid and dash lines representing median, upper, and lower quartiles, respectively.

5.2.3 Hierarchical clustering and batch effect correction

All analyses were performed using R (R Development Core Team, 2008). Hierarchical clustering was performed using function *hclust*, with 1-correlation as distance. Hierarchical clustering of log2-RPKM clearly separated samples by species, as well as RNA extraction protocols (Supplementary fig. 1a). To normalize the expression values, we performed quantile normalization to log2-RPKM values (Hansen, Irizarry et al. 2012), and eliminated batch effects using R package *ComBat* (Johnson et al., 2007). We then built a hierarchical clustering with the normalized data, which resulted in a clustering by species and regions (Supplementary fig 1b).

5.2.4 Principal component analysis

We performed a PCA for all samples included in the study, and for samples belonging to each species separately, using the function *prcomp*. The first three components were plotted using the R package *rgl* (Adler and Murdoch, 2010).. The percentage of variance explained by each component was also plotted (Figure 1).

5.2.5 Effects of RIN and PMI on mRNA expression

To investigate the effects of main factors and confounders on the data, we drew the first three principle components against factors, including species, region, RIN, and PMI.

The number of genes with RPKM > 1 and RPKM > 10 were plotted against RIN and PMI, respectively. The upper quartile, median, and lower quartile of gene RPKM were plotted against RIN and PMI. No correlation was found in all these plots. As chimpanzee PMI were not available and all macaque samples were processed within one hour after death, only human sample were used in PMI plots.

5.3 Identification of DEX genes

DEX genes were computed with the R package *DESeq* (Anders and Huber, 2010). In this package, the count data were modeled with a negative binomial distribution to address the problem of overdispersion.

5.3.1 Identification of inter-species DEX genes

Inter-species DEX genes were computed for each brain region. Five generalized linear model of the negative binomial family with log link were fitted for each gene. The fitted models were listed as follows:

- Model A: count ~ species + batch (3 levels of *species*: human, chimpanzee, macaque)
- Model H: count ~ species + batch (2 levels of *species*: human, chimpanzee/macaque)
- Model P: count ~ species + batch (2 levels of *species*: chimpanzee, human/macaque)
- Model M: count ~ species + batch (2 levels of *species*: macaque, human/chimpanzee)
- Model 0: count ~ batch

P values were calculated by a χ^2 likelihood ratio test comparing Model A with Model 0 and inter-species DEX genes were called at a false discovery rate (FDR; Benjamini and Hochberg, 1995) of 0.01.

Eleven inter-species gene expression patterns were investigated per region. The genes were called at false discovery rate 0.01 for significant differential expression and 0.1 for non-significance. The models used were listed in Supplementary Table 3.

5.3.2 Identification of intra-species DEX genes

P values of likelihood ratio test were computed by comparing the alternative model (count ~ region + batch) with the null model (count ~ batch). Intra-species DEX genes were called at a false discovery rate (FDR) of 0.01.

5.4 Weighted gene co-expression network analysis (WGCNA)

Signed co-expression networks were built using the *WGCNA* package in R (Langfelder and Horvath, 2008). To reduce noise, only inter- and intraspecies DEX genes were included in the analysis. For all genes included, a pair-wise correlation matrix was computed, and an adjacency matrix was calculated by raising the correlation matrix to a power. The power was set to 15 according to a scale-free topology criterion. For each pair of genes, a robust measure of network interconnectedness (topological overlap measure) was calculated based on the adjacency matrix. The topological overlap based dissimilarity was then used as input for average linkage hierarchical clustering. Modules were generated by hybrid dynamic tree-cutting. To obtain rare expression patterns, we set the

minimum module size to 5 genes, deepSplit to 4, and the minimum height for merging modules to 0.15.

Each module was summarized by an eigengene, which is the first principal component of the scaled module expression. Thus, the module eigengene explained the maximum amount of variation of the module expression levels. The module membership (also known as module eigengene based connectivity kME) is defined as the correlation between gene expression values and the module eigengene. To obtain modules with clear clustering of expression patterns, we iteratively reassigned genes to the module with largest kME and genes with maximum kME < 0.5 were assigned to the M0.

5.5 Conservation of neurotransmission systems

We compared networks of neurotransmitter receptors across species. Networks were constructed based on pairwise correlation of receptors' subunits genes from the same neurotransmission system. The preservation of the networks was evaluated with the statistics computed by *modulePreservation* function in *WGCNA* R package.

6. Functional enrichment analysis

Functional enrichment was assessed using DAVID Bioinformatics Resources (http://david.abcc.ncifcrf.gov; Huang et al., 2009).

7. Determination of the genetic relationships among samples

We used the DFC samples from each individual (expect for HSB 136, for which OFC was used, as the DFC sample was filtered out in the process of quality control) to study the relationships between species, under the assumption that one sample should represent the true phylogeny of the species.

7.1 Mapping

All DFC samples were mapped against the human reference genome by GEM (Marco-Sola et al., 2012), allowing mismatches based on different divergence of the species. Human samples were mapped at edit distance 1. Unmapped reads were trimmed at 50 bp and mapped at edit distance 1. Chimps were mapped at edit distance 2. Unmapped reads were then mapped at edit distance 3. If still unmapped, reads were then trimmed at 50 bp and mapped at edit distance 2. Macaque samples were mapped at edit distance 2; unmapped reads were then mapped at the edit distance 4, and 6, sequentially. If still unmapped, reads were trimmed at 50 bp and mapped at edit distance 2, 4 and 6, sequentially.

7.2 Identification of single nucleotide variants

Regions with more than 20X coverage in all the samples were selected for the SNP calling. This threshold was selected because only one sample is used for this analysis. In total, 3,373 Kbps were used. SNP calling was performed using standard SAMtools methods (Li, Ruotti et al. 2010) and SNPs with more than two variants were filtered out.

7.3 Construction of phylogentic tree

PLINK (Purcell, Neale et al. 2007) was used to calculate a distance matrix between the samples. The distance is defined as the average number of positions in which the two individuals have a different base over the total number of variable sites. MEGA software (Tamura, Peterson et al. 2011) was used to build a neighbor-joining tree from the distance matrix.

8. Validation of dopamine synthesis genes

8.1 Digital PCR

We employed Droplet Digital PCR (ddPCR) to reliably quantitate gene expression. An aliquot of the total RNA that was previously extracted from each brain sample was used for secondary validation through ddPCR analysis. One µg of total RNA was used for cDNA synthesis using SuperScript III First-strand synthesis Supermix (Invitrogen) and subsequently diluted with nuclease-free water. Custom gene-specific primers and probe for each gene of interest were designed using NCBI/Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and PrimerQuest tool (IDT). ddPCR (Hindson et al. 2011) was carried out using the Bio-Rad QX100 system. After each PCR reaction mixture consisting of ddPCR master mix and custom primers/probe set was partitioned into 15,000–20,000 droplets, parallel PCR amplification was carried out. Endpoint PCR signals were quantified and Poisson statistics was applied to yield target copy number quantification of the sample. Two-color PCR reaction was utilized for the normalization of gene expression by the housekeeping gene *TBP*. Table 6 provides sequences of primers and probes used for the validation.

8.2 Immunoblotting

We used 20 µg of proteins for immunoblotting. Blots were blocked using 5% fat-free milk in TBST for 1h, and then exposed to corresponding primary antibodies for 1h. After extensive washings, blotted protein was exposed to appropriate horseradish peroxidase–conjugated secondary antibodies and visualized with chemi-luminescence systems (Chemiglow).

9. Supplementary Tables

Supplementary Table1 | Tissue sample metadata

This table summarizes quality control data for individual tissue samples, including RNA integrity number (RIN) and RNA extracting kit used.

Supplementary Table2 | Models used to categorize interspecies DEX genes

List of *post hoc* comparisons used to categorize species-specific differential expressed genes.

Supplementary Table 3 | Human-specific DEX genes

List of DEX genes that are up- or downregulated in humans, when compared to chimpanzee and macaque.

Supplementary Table 4 | Weighted co-expression network modules

List of 229 modules corresponding to regional patterns identified by weighted gene co-expression network analysis (WGCNA); list of intra-modular genes and their intra-modular connectivity; list of up- and downregulated patterns and the species where they are found.

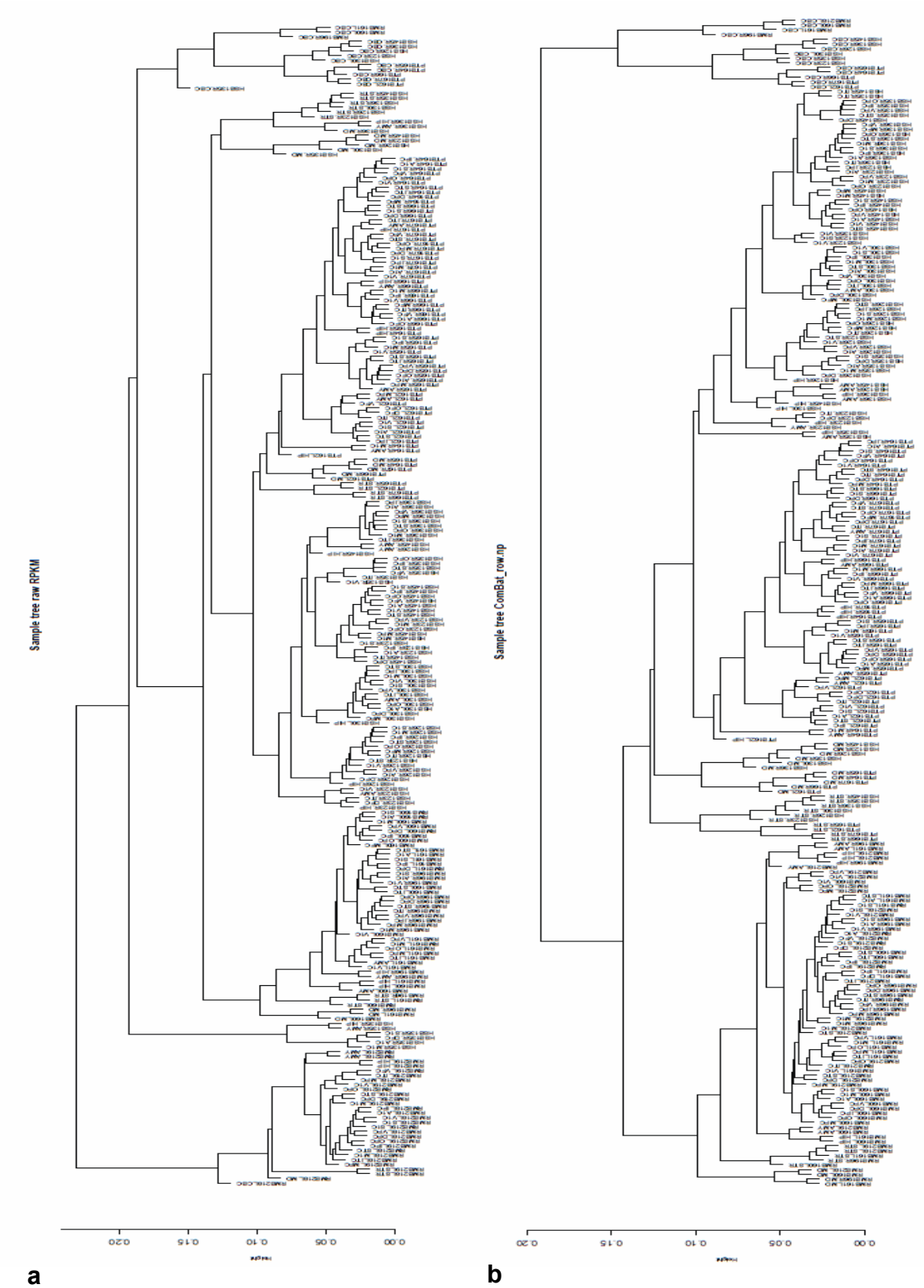
Supplementary Table 5 | List of spike-in RNAs

List of spike-in RNAs used in this study and the Barcode scheme used for identification of each brain region.

Supplementary Table 6 | List of primers used for ddPCR validation

List of primers and fluorescent oligo probes used in this study for validation of results using ddPCR. *TBP* was used as the housekeeping gene.

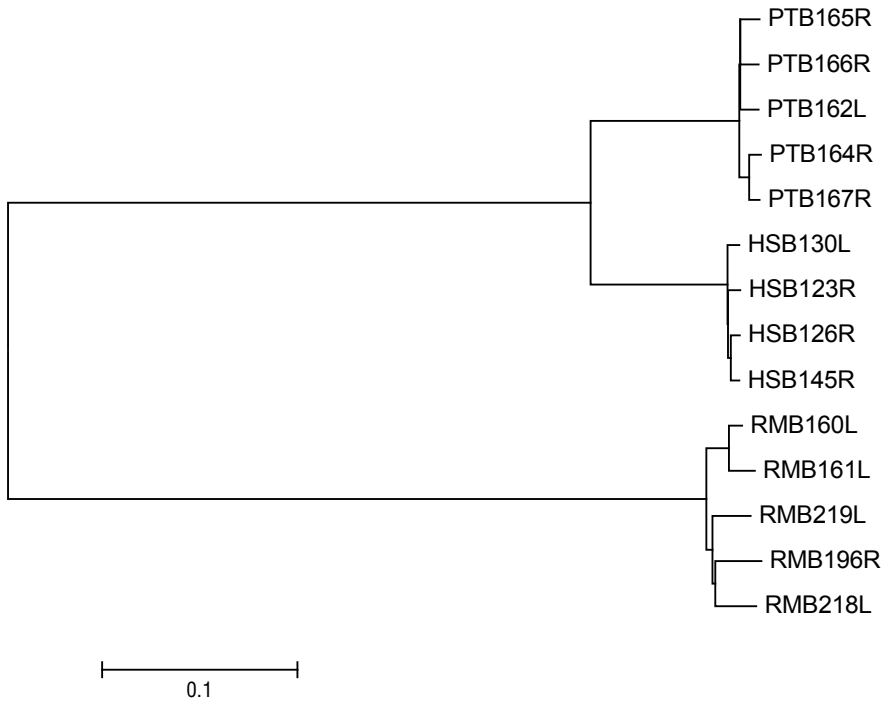
10. Supplementary Figures



a b

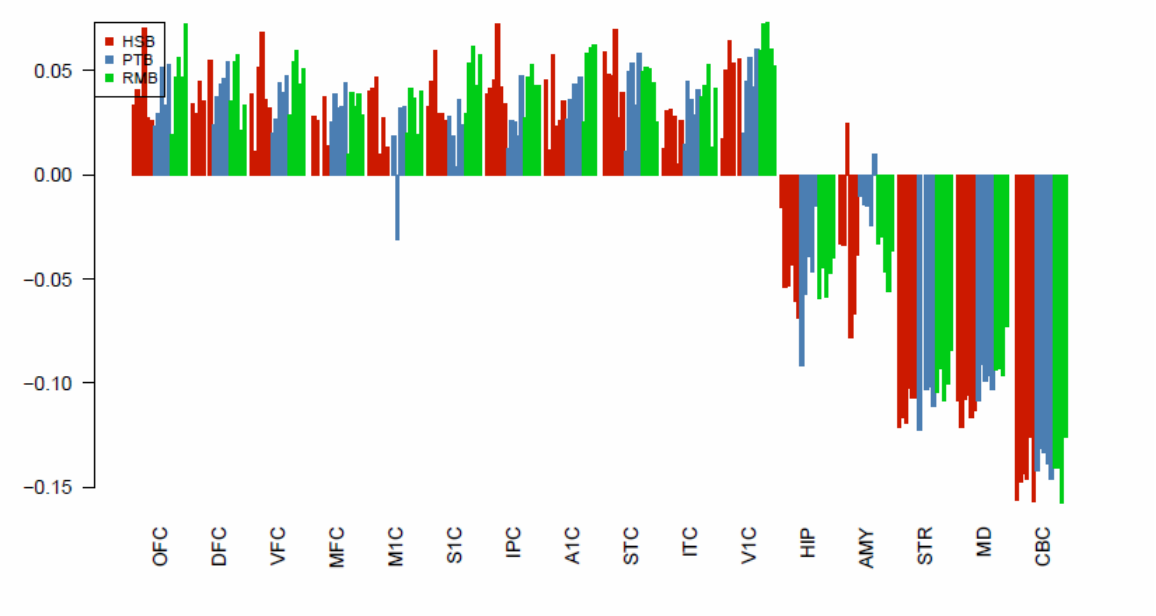
Supplementary Figure 1. Unsupervised sample clustering

Diagram showing the unsupervised hierarchical clustering of all samples used in this study before (**a**) and after batch effect correction using *ComBat* (**b**).



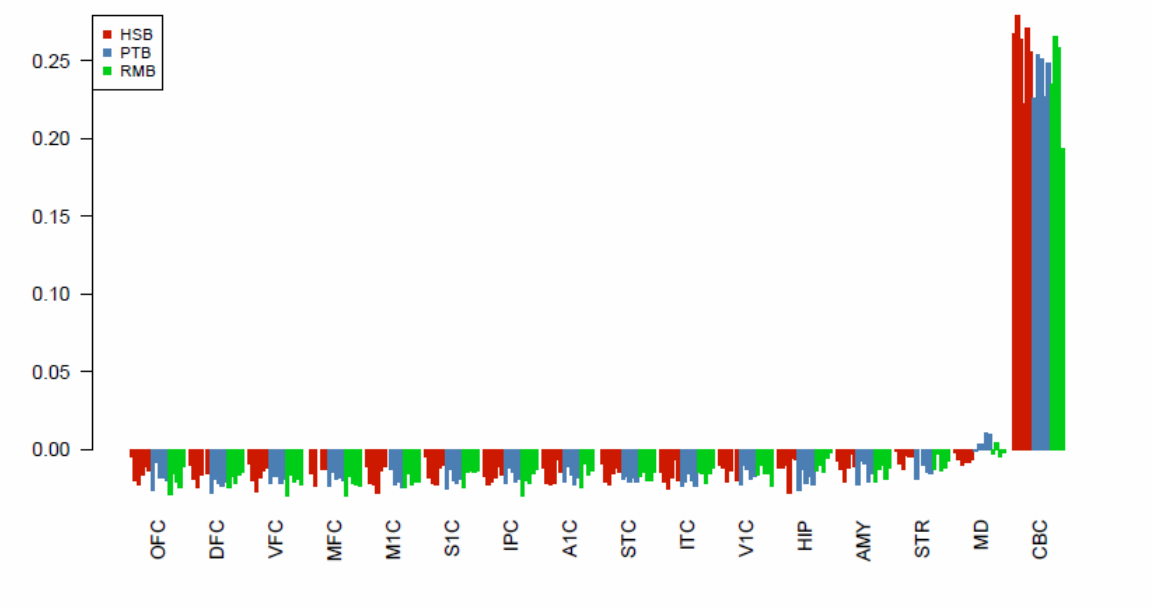
Supplementary Figure 2. Species neighbor-joining tree from variant calling

Clustering of DFCs from each specimen based on genetic distances. All samples from each species clustered together. The distance is corresponds to the average number of positions in which two individuals have a different base over the total number of variable sites.



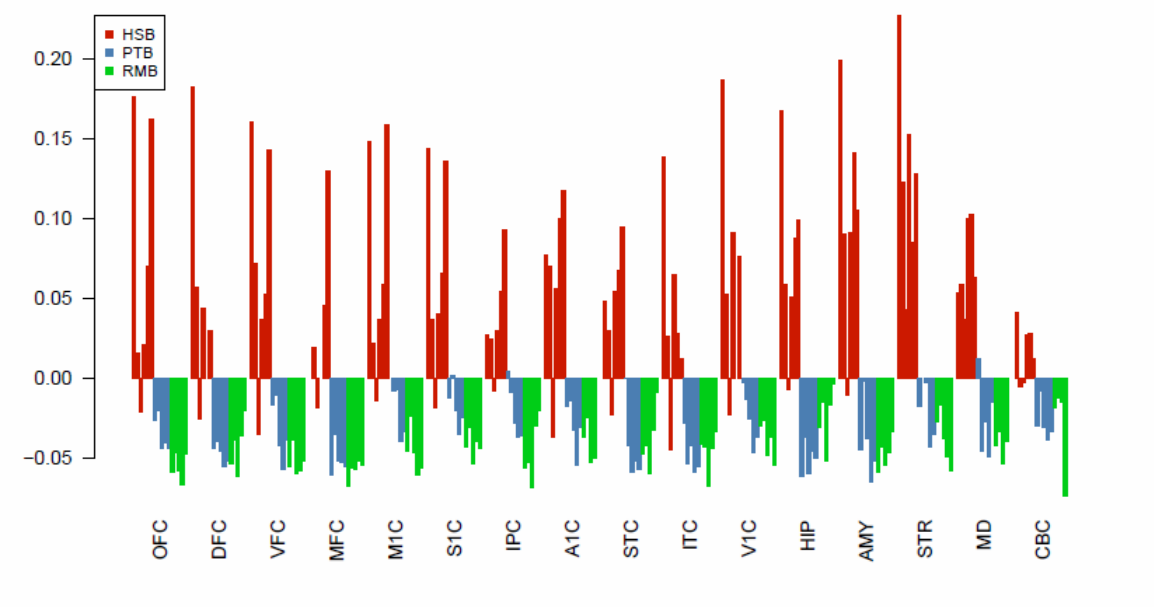
Supplementary Figure 3. WGCNA Module 33

The bar plot shows the regional pattern of the module. M33 was associated with an upregulation in the NCX areas of all species.



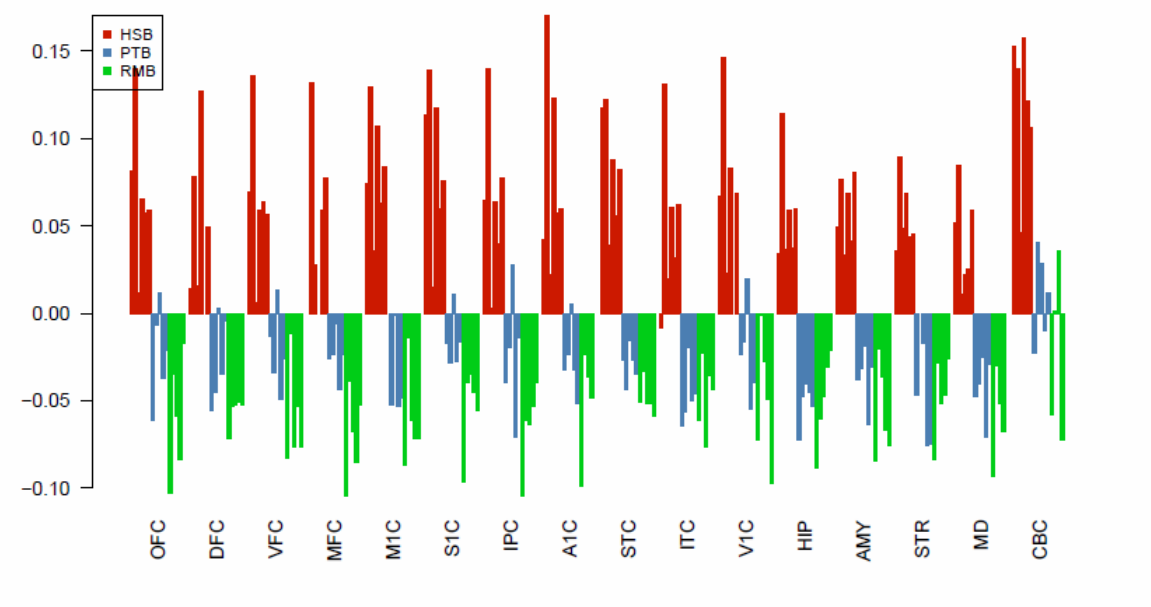
Supplementary Figure 4. WGCNA Module 16

The bar plot shows the regional pattern of the module. M16 was associated with an upregulation in the CBC areas of all species.



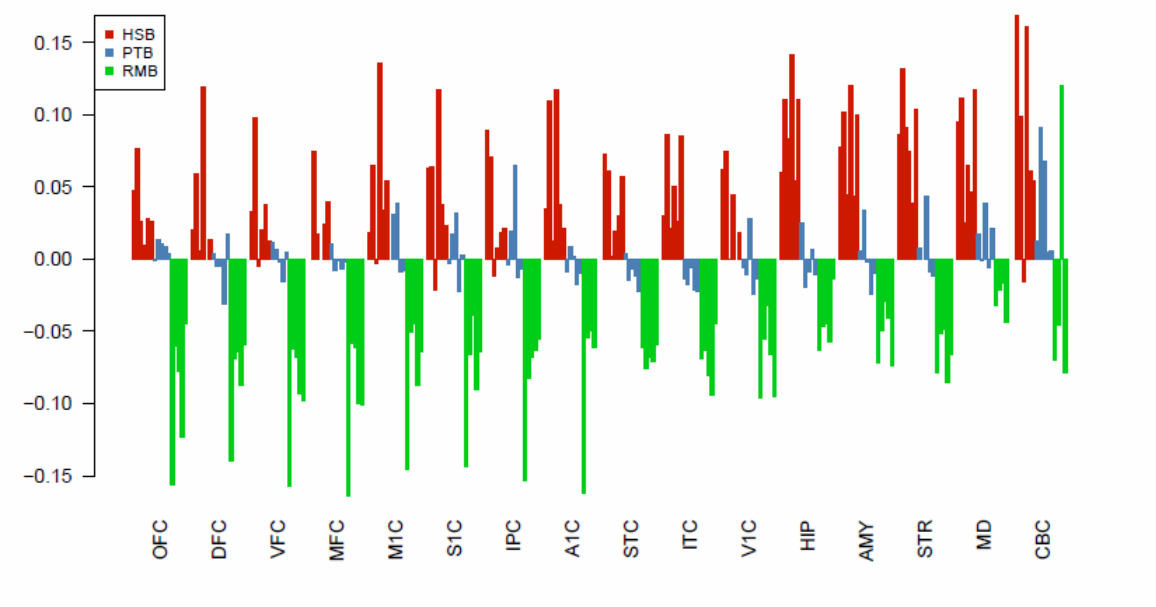
Supplementary Figure 5. WGCNA Module 81

The bar plot shows the regional pattern of the module. M81 was associated with an upregulation in all regions of the human brain.



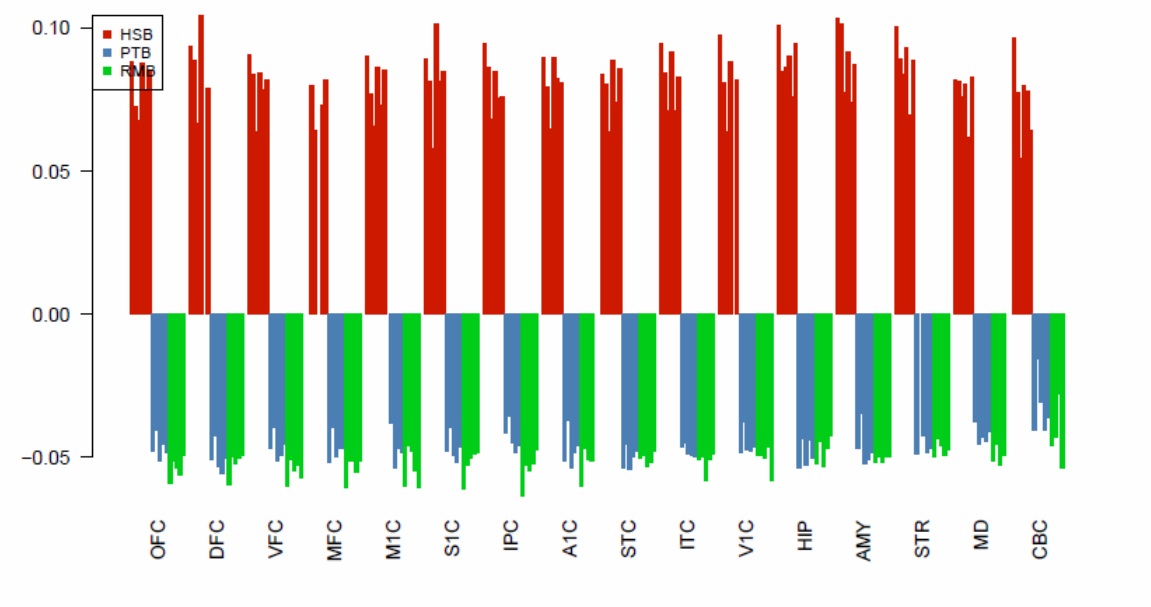
Supplementary Figure 6. WGCNA Module 162

The bar plot shows the regional pattern of the module. M162 was associated with an upregulation in all regions of the human brain.



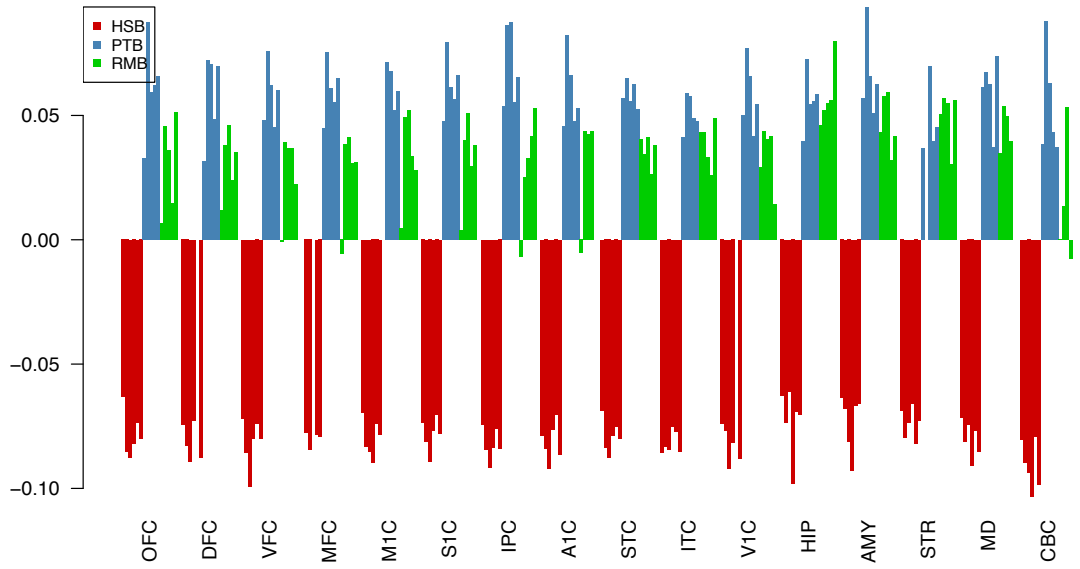
Supplementary Figure 7. WGCNA Module 192

The bar plot shows the regional pattern of the module. M192 was associated with an upregulation in all regions of the human brain.



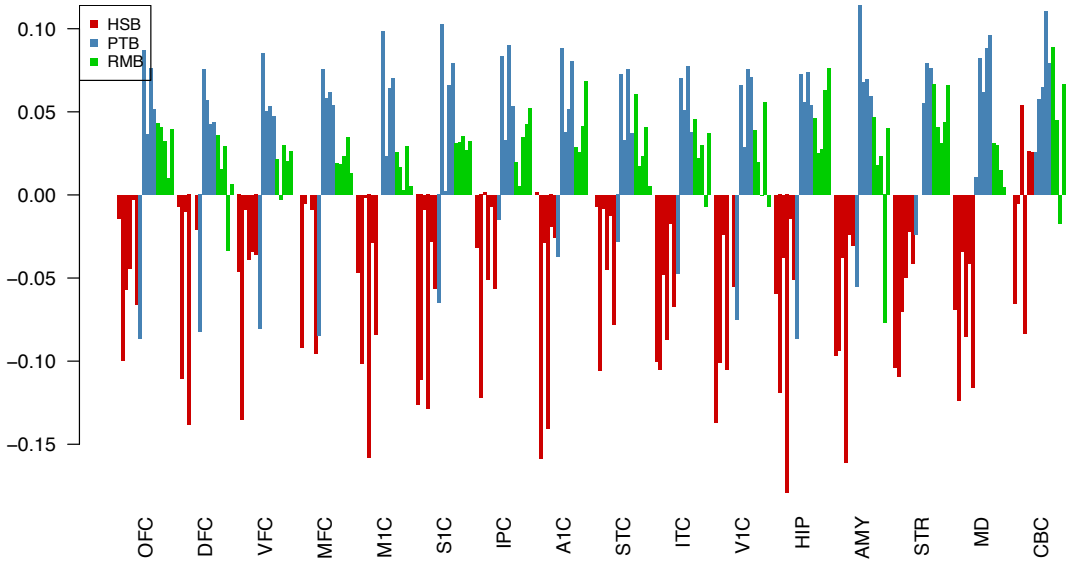
Supplementary Figure 8. WGCNA Module 229

The bar plot shows the regional pattern of the module. M229 was associated with an upregulation in all regions of the human brain.



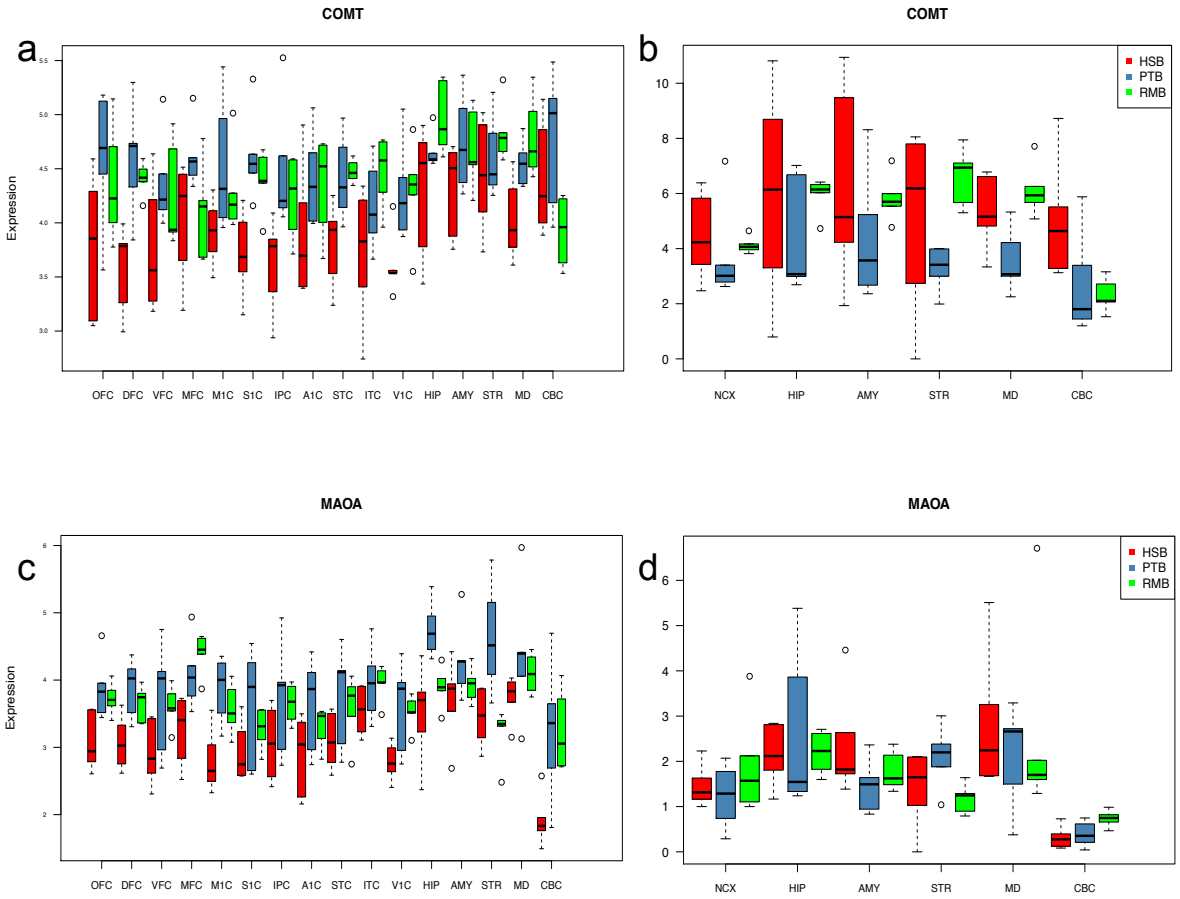
Supplementary Figure 9. WGCNA Module 69

The bar plot shows the regional pattern of the module. M69 was associated with a downregulation in all regions of the human brain.



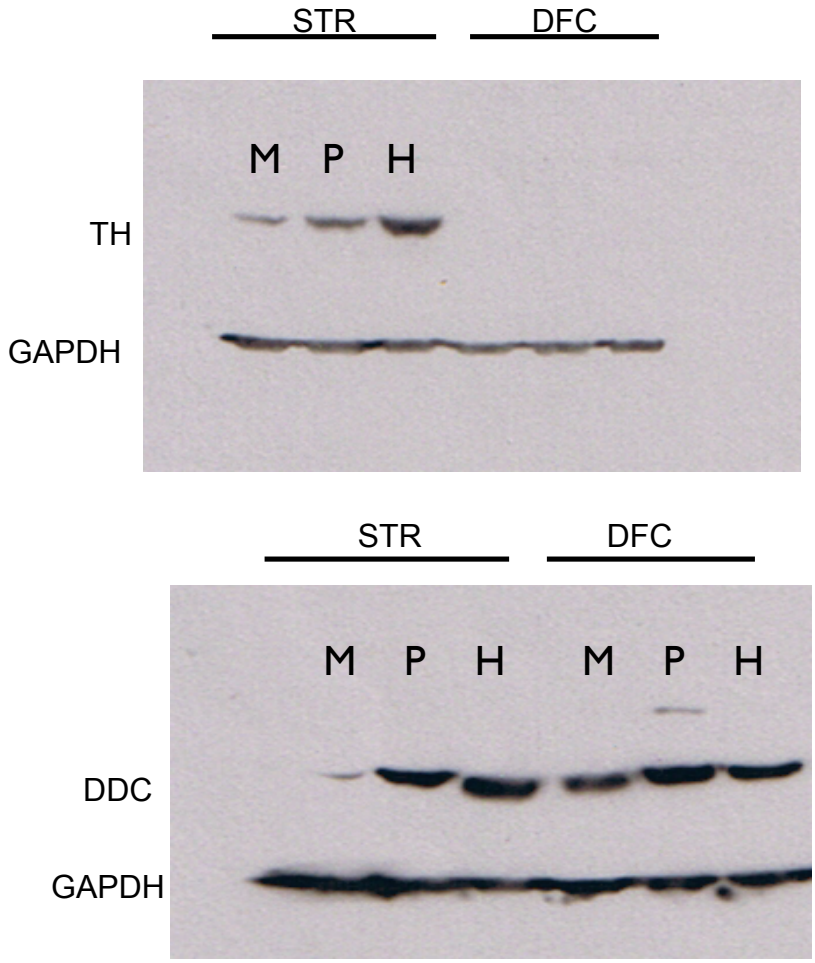
Supplementary Figure 10. WGCNA Module 173

The bar plot shows the regional pattern of the module. M173 was associated with a downregulation in all regions of the human brain.



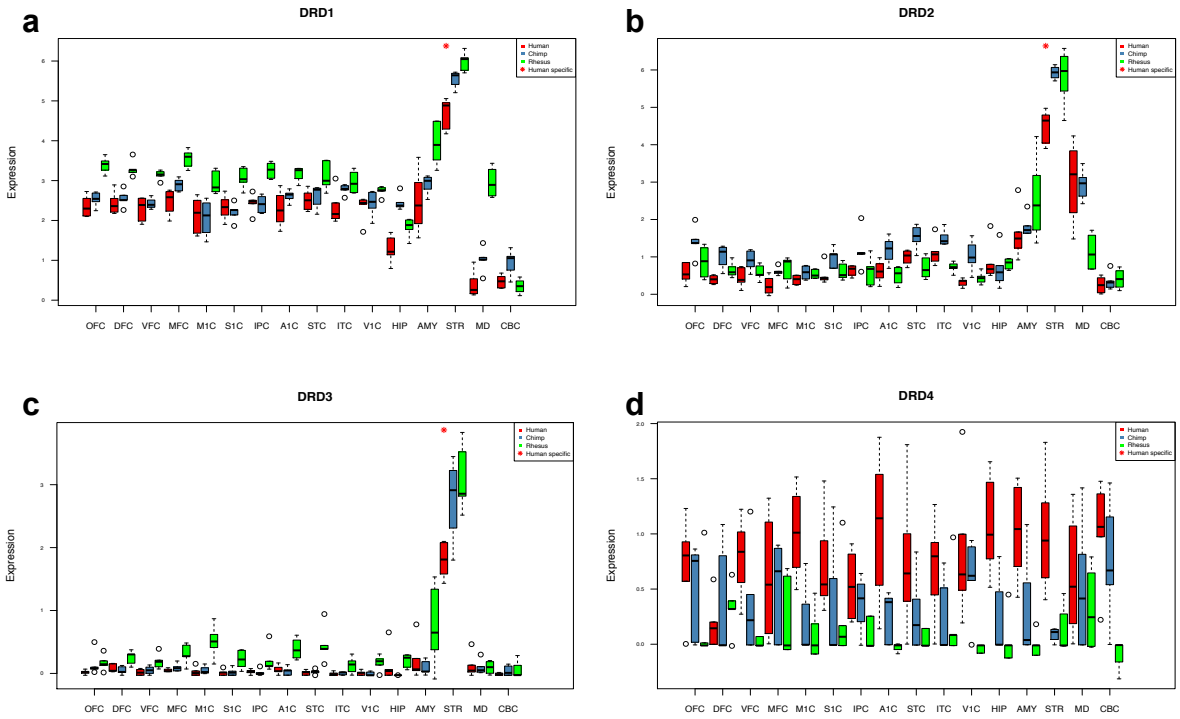
Supplementary Figure 11. Expression pattern of genes involved in dopamine degradation

a, Expression [$\log_2$ (RPKM+1)] pattern of *MAOA*. **b**, ddPCR validation of *COMT* expression normalized to housekeeping gene *TBP.* **c**, Expression pattern of *COMT.* Although human expression values of both genes are lower than in other species no significant difference was observed. **d**, ddPCR validation of *MAOA* expression normalized to housekeeping gene *TBP.*



Supplementary Figure 12. Immunoblots of TH and DDC in STR and DFC of human, chimpanzee and macaque brains.

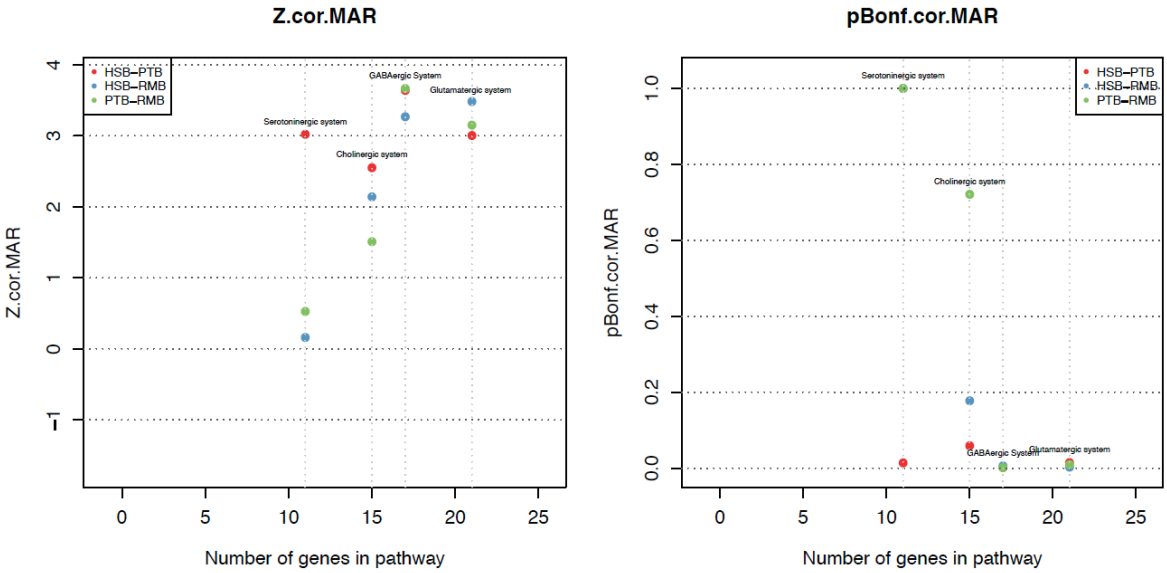
TH and DDC protein levels in STR and DFC of human (H), chimpanzee (C) and macaque (M) brains. TH is significantly more expressed in the human STR compared to the STR of chimpanzee and macaque. Similarly, DDC show lower levels in STR of macaques and higher levels in human STR.



Supplementary Figure 13. Expression pattern of genes coding dopamine receptors

Expression [$\log_2$ (RPKM+1)] pattern of **a**, *DRD1*, **b**, *DRD2*, **c**, *DRD3*, and **d**, *DRD4.*

All genes, except *DRD4*, are significantly downregulated in the human STR.



Supplementary Figure 14. Neurotransmitter networks conservation

Figure showing Z statistics of Maximum Adjacency Ratio (MAR; Langfelder et al., 2011) and corresponding P values of neurotransmitter receptors networks. Cholinergic and serotoninergic systems are less conserved, whereas glutamatergic and GABAergic are highly conserved.

3.10. References

Anders S, Huber W. (2010). “Differential expression analysis for sequence count data”. Genome Biology **11**(10):R106.

Babbitt, C. C., O. Fedrigo, A. D. Pfefferle, A. P. Boyle, J. E. Horvath, T. S. Furey and G. A. Wray (2010). "Both Noncoding and Protein-Coding RNAs Contribute to Gene Expression Evolution in the Primate Brain." Genome Biology and Evolution **2**: 67-79.

Bejerano, G., D. Haussler and M. Blanchette (2004). "Into the heart of darkness: large-scale clustering of human non-coding DNA." Bioinformatics **20**: 40-48.

Britanova, O., C. de Juan Romero, A. Cheung, K. Y. Kwan, M. Schwark, A. Gyorgy, T. Vogel, S. Akopov, M. Mitkovski, D. Agoston, N. Sestan, Z. Molnar and V. Tarabykin (2008). "Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex." Neuron **57**(3): 378-392.

Broca, P. (1861). "Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche." Bulletin de la Société d'Anthropologie **2**: 235–238.

Caceres, M., J. Lachuer, M. A. Zapala, J. C. Redmond, L. Kudo, D. H. Geschwind, D. J. Lockhart, T. M. Preuss and C. Barlow (2003). "Elevated gene expression levels distinguish human from non-human primate brains." Proceedings of the National Academy of Sciences of the United States of America **100**(22): 13030-13035.

Colantuoni, C., B. K. Lipska, T. Z. Ye, T. M. Hyde, R. Tao, J. T. Leek, E. A. Colantuoni, A. G. Elkahloun, M. M. Herman, D. R. Weinberger and J. E. Kleinman (2011). "Temporal dynamics and genetic control of transcription in the human prefrontal cortex." Nature **478**(7370): 519-U117.

Daniel Adler and Duncan Murdoch (2010). “rgl: 3D visualization device system (OpenGL)”. R package, version 0.92.794. http://CRAN.R-project.org/package=rgl

Dax, M. (1865). "Lésions de la moitié gauche de l'encéphale coincidant avec l'oubli des signes de la pensée (lu au Congrés méridional tenu à Montpellier en 1836)." Gaz Heb Méd Chi.

de Lucchi, R., M., B. J. Dennis and W. R. Adey (1965). Stereotaxic Atlas of the Chimpanzee Brain (Pan satyrus). , University of California Press.

Enard, W., P. Khaitovich, J. Klose, S. Zollner, F. Heissig, P. Giavalisco, K. Nieselt-Struwe, E. Muchmore, A. Varki, R. Ravid, G. M. Doxiadis, R. E. Bontrop and S. Paabo (2002). "Intra- and interspecific variation in primate gene expression patterns." Science **296**(5566): 340-343.

Gu, J. Y. and X. Gu (2003). "Induced gene expression in human brain after the split from chimpanzee." Trends in Genetics **19**(2): 63-65.

Habegger, L., A. Sboner, T. A. Gianoulis, J. Rozowsky, A. Agarwal, M. Snyder and M. Gerstein (2011). "RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries." Bioinformatics **27**(2): 281-283.

Han, W., K. Y. Kwan, S. Shim, M. M. Lam, Y. Shin, X. Xu, Y. Zhu, M. Li and N. Sestan (2011). "TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract." Proceedings of the National Academy of Sciences of the United States of America **108**(7): 3041-3046.

Hansen, K. D., R. A. Irizarry and Z. J. Wu (2012). "Removing technical variability in RNA-seq data using conditional quantile normalization." Biostatistics **13**(2): 204-216.

Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Chrast J, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigó R, Hubbard TJ. (2012). “GENCODE: the reference human genome annotation for The ENCODE Project”. Genome Research **22**(9):1760-1774.

Hevner, R., L. Shi, N. Justice, Y. Hsueh, M. Sheng, S. Smiga, A. Bulfone, A. Goffinet, A. Campagnoni and J. Rubenstein (2001). "Tbr1 regulates differentiation of the preplate and layer 6." Neuron **29**(2): 353-366.

Hindson BJ, Ness, KD, Masquelier, DA, Belgrader, P, Heredia, NJ, Makarewicz, AJ, Bright, IJ, Lucero, MY, Hiddesen, AL, Legler, TC, Kitano, TK, Hodel, MR, Petersen, JF, Wyatt, PW, Steenblock, ER, Shah, PH, Bousse, LJ, Troup, CB, Mellen, JC, Wittman, DK, Erndt, NG, Cauley, TH, Koehler, RT, So, AP, Dube, S, Rose, KA, Montesclaros, L, Wang, S, Stumbo, DP, Hodges, SP, Romine, S, Milanovich, FP, White, HE, Regan, JF, Karlin-Neuman, GA, Hindson, CM, Saxonov, S, Colston, BW. (2011). High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Analytical Chemistry **83**(22): 8604–8610.

Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, Hillman-Jackson J, Kuhn RM, Pedersen JS, Pohl A, Raney BJ, Rosenbloom KR, Siepel A, Smith KE, Sugnet CW, Sultan-Qurraie A, Thomas DJ, Trumbower H, Weber RJ, Weirauch M, Zweig AS, Haussler D, Kent WJ. (2006). “The UCSC Genome Browser Database: update 2006”. Nucleic Acids Res **34**: D590–D598.

Huang, D. W., B. T. Sherman and R. A. Lempicki (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nature Protocols **4**(1): 44-57.

Johnson, M. B., Y. I. Kawasawa, C. E. Mason, Z. Krsnik, G. Coppola, D. Bogdanovic, D. H. Geschwind, S. M. Mane, M. W. State and N. Sestan (2009). "Functional and evolutionary insights into human brain development through global transcriptome analysis." Neuron **62**(4): 494-509.

Johnson, WE, Rabinovic, A, and Li, C (2007). Adjusting batch effects in microarray expression data using Empirical Bayes methods. Biostatistics **8**(1):118-127.

Kang, H. J., Y. I. Kawasawa, F. Cheng, Y. Zhu, X. Xu, M. Li, A. M. Sousa, M. Pletikos, K. A. Meyer, G. Sedmak, T. Guennel, Y. Shin, M. B. Johnson, Z. Krsnik, S. Mayer, S. Fertuzinhos, S. Umlauf, S. N. Lisgo, A. Vortmeyer, D. R. Weinberger, S. Mane, T. M. Hyde, A. Huttner, M. Reimers, J. E. Kleinman and N. Sestan (2011). "Spatio-temporal transcriptome of the human brain." Nature **478**(7370): 483-489.

Kent WJ. (2002). “BLAT--the BLAST-like alignment tool”. Genome Research: **12**(4):656-64.

Khaitovich, P., B. Muetzel, X. W. She, M. Lachmann, I. Hellmann, J. Dietzsch, S. Steigele, H. H. Do, G. Weiss, W. Enard, F. Heissig, T. Arendt, K. Nieselt-Struwe, E. E. Eichler and S. Paabo (2004). "Regional patterns of gene expression in human and chimpanzee brains." Genome Research **14**(8): 1462-1473.

King, M. and A. Wilson (1975). "Evolution at two levels in humans and chimpanzees." Science **188**(4184): 107-116.

Langfelder, P. and S. Horvath (2008). "WGCNA: an R package for weighted correlation network analysis." BMC Bioinformatics **9**: 559.

Langfelder, P., Luo, R., Oldham, M. D., and S. Horvath (2011). "Is my network module preserved and reproducible?" PLoS Computational Biology **7**: e1001057.

Li, B., V. Ruotti, R. M. Stewart, J. A. Thomson and C. N. Dewey (2010). "RNA-Seq gene expression estimation with read mapping uncertainty." Bioinformatics **26**(4): 493-500.

Liu, X. L., M. Somel, L. Tang, Z. Yan, X. Jiang, S. Guo, Y. Yuan, L. He, A. Oleksiak, Y. Zhang, N. Li, Y. H. Hu, W. Chen, Z. L. Qiu, S. Paabo and P. Khaitovich (2012). "Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques." Genome Research **22**(4): 611-622.

Lockhart, D. J. and C. Barlow (2001). "Expressing what's on your mind: DNA arrays and the brain." Nature Reviews Neuroscience **2**(1): 63-68.

Marco-Sola, S, Sammeth, M, Guigó, R, Ribeca, P. (2012). “The GEM mapper: fast, accurate and versatile alignment by filtration”. Nature Methods **9**: 1185–1188.

McKenna, W. L., J. Betancourt, K. A. Larkin, B. Abrams, C. Guo, J. L. Rubenstein and B. Chen (2011). "Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development." J Neurosci **31**(2): 549-564.

Mikkelsen, T. S., L. W. Hillier, E. E. Eichler, M. C. Zody, D. B. Jaffe, S. P. Yang, W. Enard, I. Hellmann, K. Lindblad-Toh, T. K. Altheide, N. Archidiacono, P. Bork, J. Butler, J. L. Chang, Z. Cheng, A. T. Chinwalla, P. deJong, K. D. Delehaunty, C. C. Fronick, L. L. Fulton, Y. Gilad, G. Glusman, S. Gnerre, T. A. Graves, T. Hayakawa, K. E. Hayden, X. Q. Huang, H. K. Ji, W. J. Kent, M. C. King, E. J. Kulbokas, M. K. Lee, G. Liu, C. Lopez-Otin,

K. D. Makova, O. Man, E. R. Mardis, E. Mauceli, T. L. Miner, W. E. Nash, J. O. Nelson, S. Paabo, N. J. Patterson, C. S. Pohl, K. S. Pollard, K. Prufer, X. S. Puente, D. Reich, M. Rocchi, K. Rosenbloom, M. Ruvolo, D. J. Richter, S. F. Schaffner, A. F. A. Smit, S. M. Mortazavi, A., B. A. Williams, K. Mccue, L. Schaeffer and B. Wold (2008). "Mapping and quantifying mammalian transcriptomes by RNA-Seq." Nature Methods **5**(7): 621-628.

Smith, M. Suyama, J. Taylor, D. Torrents, E. Tuzun, A. Varki, G. Velasco, M. Ventura, J. W. Wallis, M. C. Wendl, R. K. Wilson, E. S. Lander, R. H. Waterston and C. S. A. Consortium (2005). "Initial sequence of the chimpanzee genome and comparison with the human genome." Nature **437**(7055): 69-87.

Pollard, K. S., S. R. Salama, B. King, A. D. Kern, T. Dreszer, S. Katzman, A. Siepel, J. S. Pedersen, G. Bejerano, R. Baertsch, K. R. Rosenbloom, J. Kent and D. Haussler (2006). "Forces shaping the fastest evolving regions in the human genome." Plos Genetics **2**(10): 1599-1611.

Prabhakar, S., J. P. Noonan, S. Paabo and E. M. Rubin (2006). "Accelerated evolution of conserved noncoding sequences in humans." Science **314**(5800): 786-786.
Roth, R. B., P. Hevezi, J. Lee, D. Willhite, S. M. Lechner, A. C. Foster and A. Zlotnik (2006). "Gene expression analyses reveal molecular relationships among 20 regions of the human CNS." Neurogenetics **7**(2): 67-80.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly and P. C. Sham (2007). "PLINK: A tool set for whole-genome association and population-based linkage analyses." American Journal of Human Genetics **81**(3): 559-575.

R Development Core Team (2009). “R: A language and environment for statistical computing”. Vienna, R Foundation for Statistical Computing.

Saleem, K. S. and N. Logothetis (2007). A combined MRI and histology atlas of the rhesus monkey brain in stereotaxic coordinates. London ; Burlington, MA, Academic.

Sequencing, R. M. G., A. Consortium, et al., (2007). "Evolutionary and Biomedical Insights from the Rhesus Macaque Genome." Science **316**(5822): 222-234.

Somel, M., H. Franz, Z. Yan, A. Lorenc, S. Guo, T. Giger, J. Kelso, B. Nickel, M. Dannemann, S. Bahn, M. J. Webster, C. S. Weickert, M. Lachmann, S. Paabo and P. Khaitovich (2009). "Transcriptional neoteny in the human brain." Proceedings of the National Academy of Sciences of the United States of America **106**(14): 5743-5748.

Srinivasan, K., D. P. Leone, R. K. Bateson, G. Dobreva, Y. Kohwi, T. Kohwi-Shigematsu, R. Grosschedl and S. K. McConnell (2012). "A network of genetic repression and derepression specifies projection fates in the developing neocortex." Proceedings of the National Academy of Sciences of the United States of America **109**(47): 19071-19078.

Stedman, H. H., B. W. Kozyak, A. Nelson, D. M. Thesier, L. T. Su, D. W. Low, C. R. Bridges, J. B. Shrager, N. Minugh-Purvis and M. A. Mitchell (2004). "Myosin gene mutation correlates with anatomical changes in the human lineage." Nature **428**(6981): 415-418.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar (2011). "MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods." Molecular Biology and Evolution **28**(10): 2731-2739.

Uddin, M., D. E. Wildman, G. Z. Liu, W. B. Xu, R. M. Johnson, P. R. Hof, G. Kapatos, L. I. Grossman and M. Goodman (2004). "Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles." Proceedings of the National Academy of Sciences of the United States of America **101**(9): 2957-2962.

Vargha-Khadem, F., D. G. Gadian, A. Copp and M. Mishkin (2005). "FOXP2 and the neuroanatomy of speech and language." Nature Reviews Neuroscience **6**(2): 131-138.

Visel, A., S. Prabhakar, J. A. Akiyama, M. Shoukry, K. D. Lewis, A. Holt, I. Plajzer-Frick, V. Afzal, E. M. Rubin and L. A. Pennacchio (2008). "Ultraconservation identifies a small subset of extremely constrained developmental enhancers." Nature Genetics **40**(2): 158-160.

Wang, Y. Q. and B. Su (2004). "Molecular evolution of microcephalin, a gene determining human brain size." Human Molecular Genetics **13**(11): 1131-1137.

Wernicke, C. (1874). Der Aphasische Symptomenkomplex: Eine Psychologische Studie auf Anatomischer Basis. Breslau, Max Cohn & Weigert.

Zhang, B. and S. Horvath (2005). "A general framework for weighted gene co-expression network analysis." Stat Appl Genet Mol Biol **4**: Article17.

CHAPTER III

GENERAL DISCUSSION, CONCLUSIONS AND FUTURE DIRECTIONS

The generation and genome-wide exon-level analysis of a comprehensive transcriptome and genotyping dataset covering multiple brain regions and NCX areas across all periods of human brain development and adulthood expands the current knowledge of the transcriptional processes that govern human brain development and provides immediate opportunities for a variety of further investigations. Also, the comparative analyses of human with chimpanzee and rhesus macaque brain transcriptomes increase our knowledge about human-specific changes that may have contributed to human distinctive features, such as language and cognition. Data from these studies are publicly available via multiple avenues, including a searchable database (www.humanbraintranscriptome.org). These data will also provide opportunities for comparison with other existing and forthcoming large-scale gene expression and transcriptome datasets, including some generated from different human conditions and model organisms (Zapala et al., 2005; Semeralul et al., 2006; Mirnics et al., 2000; Ryan et al., 2006).

Our initial analyses revealed several important aspects of the spatio-temporal dynamics and functional organization of the human brain transcriptome and provided insights into human neurobiological processes. These findings build substantially on previous finding that gene transcription and exon usage have complex and dynamically regulated patterns across space and time in the brain (Chikaraishi, 1979; Sutcliffe, 1988). We showed that global transcriptome dynamics and organization are age-dependent and differ across human brain regions. These transcriptional differences are far more prominent than those that exist between sexes, ethnicities, or individuals, in spite of their underlying genomic differences. We found that approximately 86% of genes are expressed in the sampled regions, and over 90% of these are differentially regulated, predominantly in the form of gene- and exon-level temporal regulation.

We found that there is a progressive increase in correlation of the transcriptome between anatomically and functionally related structures (e.g., NCX, HIP, AMY) during development. Furthermore, our analysis of co-expressed gene networks identified large modules comprised of genes that exhibit opposite trajectories and a reversal in gene expression across multiple regions just before birth (late fetal period). We speculate that these findings likely reflect several underlying phenomena and features of human brain development and structure. One such feature is a general trend of pronounced transcriptional differences during embryonic and fetal development, when specific neural

circuits are being established, followed by an increased correlation between functionally related regions and NCX areas in the postnatal and adult brain. Furthermore, perinatal reversals in expression of a large number of genes may reflect intrinsic changes in developmental processes between fetal and postnatal development as well as effects of early experience and environment on brain development. We also showed that the identified network modules are related to specific biological pathways and contain hub genes that encode transcriptional regulators, modulators of chromatin state, and signal transduction proteins, all of which likely play an important role in driving the module expression pattern.

We also demonstrated how this dataset can be used to identify novel spatial and temporal patterns of alternative exon usage. Specifically, we showed that an alternative isoform of *ANKRD32* forms a transient graded expression pattern in the PFC and motor cortex during midfetal development. Since the anterior-posterior gradient of *ANKRD32b* expression in human period 6 frontal cortex was not detected in mice of the equivalent developmental period, it is possible that this isoform is involved in the development and evolution of species differences in the frontal cortex. The great majority of genes and putative alternative isoforms identified in this study have not previously been characterized, emphasizing how little is known about genetic mechanisms in the developing human brain.

The accuracy and usefulness of the data was illustrated by profiling spatial differences and developmental trajectories of genes related to specific neurobiological processes and brain disorders, many of which would not likely be evident in transcriptome data generated from model organisms. Specifically, the examples presented here demonstrate how these transcriptome data can be used to assess the onset and the spatiotemporal trajectories of major neurodevelopmental processes, cell types, and neurotransmitter systems in the human brain. Notable regional differences were observed between developing brain regions for the three major categories of neurobiological processes. For example, we observed specific transcriptional changes, including peaking of *PVALB* and *UTS2D* expression, in the NCX and AMY, during adolescence, a developmental period during which the symptoms of schizophrenia and other neuropsychiatric disorders typically arise (Insel, 2010; State, 2010; Crow, 2008; Meyer-Lindenberg and Weinberger, 2006; Lewis and Levitt, 2002). This is consistent with previous observations of alterations in PVALB-immunopositive interneurons in patients with schizophrenia and seizures (Lewis and Levitt, 2002).

Our analysis confirmed and expanded on previous findings on sexually dimorphic gene expression in the human brain. Overall, more sex-specific DEX genes were found in males and during prenatal and, to a lesser extent, postnatal development than adult life. Many sex-specific genes also showed significant regional differences and are associated

with specific biological processes and major disorders. We featured two examples: *PCDH11Y*, a gene previously associated with cerebral asymmetry (Priddle and Crow, 2003), and *IGF2*, an imprinted gene implicated in embryonic growth and cognitive function (Lehtinen et al., 2011; Chen et al., 2011). The imprinting pattern of *IGF2* undergoes complex regulation in the developing mouse brain (Gregg et al., 2010), and the dramatic changes we observed may reflect underlying epigenetic regulation.

From the genes displaying DEU between males and females, we highlighted *NLGN4X*, a gene encoding a protein involved in synapse formation and function (Sudhof et al., 2008), which is expressed at comparable levels in male and female brain but displays sex-specific DEU pattern. Consistent with its role in synapse formation, *NLGN4X* is upregulated during the second half of the prenatal NCX development that coincides with an increase in genes associated with synapse and dendrite development. Mutations in *NLGN4X* have previously been associated with ASD (Jamain et al., 2003), which are more prevalent in males. Many of the ASD associated risk SNPs and deletions are present in *NLGN4X's* exons 5, 6, and 7, which display male-specific bias in DEU. While we do not yet understand the full functional implications of the findings of sex-specific gene expression and exon usage, they offer possible transcriptional mechanisms underlying the sex differences in the incidence, prevalence, and severity of many neurological and psychiatric disorders.

In addition, our data reveal previously unknown spatio-temporal expression patterns for many genes previously associated with ASD, schizophrenia, and other major disorders. Coupled with analysis of co-expressed genes in the dataset, these provide information about when and where these disease-associated genes are expressed in the developing and adult brain, which can be used to help to infer their function. Moreover, genome-wide association or linkage studies of complex multigenic disorders frequently implicate large genomic intervals that encompass dozens of relatively unknown candidate genes. Our data can enhance such results by narrowing the focus to those genes that are specifically expressed during development or restricted to a specific region known to be preferentially affected.

We show associations between specific SNPs and expression levels at both gene and exon level in different regions of the developing human brain, indicating that genetic variations contribute to inter-individual transcriptome variability across regions and development. While the current number of specimens in our dataset restricted our power to detect eQTLs, the *cis*-eQTLs that we were able to detect may provide insights into how SNPs may affect the regulation of gene expression.

In our second analyses of the human brain transcriptome, we focused on the following prominent features of the human neocortex: it is parcelated into many

functionally distinct areas; some of its cognitive and motor functions are lateralized between the two cerebral hemispheres; and its development is very protracted over prenatal and postnatal development. These three aspects of the human neocortical development remained poorly understood, mainly because until now it has been difficult to directly study them in the developing human brain.

The main finding of this analysis is the three main phases that the human neocortical transcriptome undergoes during the lifespan. The first phase corresponds to the intrauterine period, when the foundations of neocortex are being built. During that period, transcriptional differences between the areas are pronounced and may be related with areal specification (O'Leary & Sahara, 2008; O'Leary et al., 2007; Rakic, 1988). The next period lasts from birth to adolescence. At the very beginning of the period, the transcriptome abruptly becomes nearly homogenous across neocortical areas. While the foundation for the structure and connectivity are established during prenatal periods, the fine-tuning of cortical connectivity occurs postnatally (Stiles et al., 2011; Johnson et al., 2001). For that, the interaction with the outside world is of the outmost importance, and the first years of life are the first opportunity for that to happen. The increased similarity of areas during infancy and childhood represents a major finding that can influence multiple fields of medicine and psychology. Those periods are critical in cognitive development and learning (Knudsen, 2004; Jolles and Crone, 2012; Johnson et al., 2001), and the profile of the transcriptome could be an important factor in these complex processes. Moreover, during these periods of homogeneity of the transcriptome the human neocortex is more highly plastic, more prone to circumvent the functional defects due to the lesion than the adult neocortex (Anderson et al., 2011; Staudt, 2010). This described homogeneity of the transcriptome between different areas lasts until the next key point of the development: the adolescence, when we see the transition to the adult-like highly specialized profile of the neocortical areas.

In addition to discovering changes in the landscape of the transcriptome during development, we report another surprising finding. Our data shows that on the level of the transcriptome, interhemispherical differences are absent. This is perhaps the most unexpected finding. Despite our intense interest and efforts to identify asymmetry in gene expression in human brain, no differences were found, in contrast to a previously study that had fewer samples (Sun et al. 2005). This finding indicates that brain lateralization may be driven by stochastic, hormonal, epigenetic, functional connectivity-based, and/or activity-dependent mechanisms that are not reflected in gene expression. Furthermore, in accordance to our findings, a recent large-scale MRI study also failed to detect robust asymmetries on the population level (Chen et al, 2012).

Finally, the comparative analysis of adult human, chimpanzee, and macaque brains revealed several differences among species, including some that are human-specific. We found that several genes are differentially expressed among species. Focusing our analysis on human-specific DEX, we found that the striatum is the region where the highest number of genes is differentially expressed. Interestingly, most of the DEX genes in humans exhibited upregulation, a pattern not seen in chimpanzee or macaques. This result corroborates previous findings that show an enrichment of upregulated gene expression in the human brain (Cáceres et al., 2003; Gu and Gu, 2003; Khaitovich et al., 2004). Among the human-specific genes, we found three to be exclusively DEX in the human neocortex: *TWIST1, RP11-364P22.1, and CTB-78F1.1*. *TWIST1* is a transcription factor that inhibits trans-activation by MEF2 (Hamamori et al., 1997). According to a recent study (Liu et al., 2012), *MEF2A*, a member of the MEF2 family, is likely the regulator of a module of co-expressed genes involved in synaptogenesis. This study was done in the developing postnatal brain and thus it is hard to integrate the findings with ours, but the downregulation of *TWIST1* in adulthood might allow an easier synaptic reorganization in humans. Intraspecies clustering of regions based on DEX genes revealed that there is a human-specific clustering of perisylvian regions involved in speech and language. Interestingly, the human neocortex appears to be more transcriptionally similar than the neocortices from chimpanzees and macaques.

We also found several human-specific modules of co-expressed genes, including two that had genes coding for dopamine biosynthesis pathway as hub genes. The upregulation of these genes in the human striatum was confirmed with ddPCR. Immunobloting confirmed that protein levels for these genes corresponded to the gene expression levels. Since the genes involved in the anabolic pathway were upregulated and the genes involved in the catabolic pathway did not show any statistically different expression, we hypothesize that humans may have higher levels of dopamine than the other species analyzed. It has been reported that dopamine is involved in several cognitive processes, especially working memory, reasoning, and overall intelligence (Goldman-Rakic, 1998; Nieoullon, 2002. The putative higher levels of dopamine production we report may give some hints on the evolution of human cognition. However, we also observed an increase in dopamine production that was accompanied by a downregulation of genes coding the dopamine receptors D1, D2, and D3. It will be important to analyze the protein levels of these receptors to confirm that humans have lower amounts of these receptors in the STR. If confirmed, this downregulation of dopamine receptors might balance the upregulation of dopamine biosynthesis. Nevertheless, further studies on this balance between neurotransmitter and their receptors will be needed.

We extended our study to other neurotransmitters through the analysis of neurotransmitter receptor systems network correlations. We found that both the glutamatergic and the GABAergic systems were well conserved among species, whereas the serotoninergic and cholinergic systems were not. The role of serotonin in cognition (Meneses and Hong, 1995; Steckler and Sahgal, 1995) and its involvement in some neurological diseases, such as schizophrenia (Roth et al., 2004; Naughton et al., 2000) are well reported. Similarly, acetylcholine has been implicated in the processes of learning and memory (Harder et al., 1998; Sarter and Parikh, 2005) and neurodegenerative diseases (Mega, 2000). We reveal that the organization of the correlation networks for these neuromodulators show some differences among humans, chimpanzees, and macaques. Interestingly, two recent studies (Raghanti et al., 2008a; Raghanti et al., 2008b) showed that there are differences at the level of prefrontal cortex innervation among these species, reinforcing that these two neuromodulatory systems might be relevant to species differences in cognition.

Together, these data highlight the complexity of spatio-temporal gene expression and exon usage, as well as the species-differences that may have been crucial for some aspects of human-specific brain function. However, there are several potential limitations in our data that warrant discussion. Foremost, we used stringent criteria in order to faithfully characterized general transcriptional patterns and minimize false positives, rather than to capture all the changes that we think may in fact occur. We also analyzed dissected tissue samples that contain multiple cell types, thus diluting the contribution and dynamic range of genes expressed in a more cell type specific manner. Current limitations prevent us from using cell type-specific approaches in systematically analyzing the transcriptome across multiple regions and all developmental periods. Also, all of these analyses were performed in postmortem tissue. Even though some studies showed that there is no correlation of PMI with RNA degradation (Barton et al., 1993; Heinrich et al., 2007), others suggested that there is a PMI-related RNA degradation that might cause a significant impact on the RNA population structure (Catts et al., 2005). Furthermore, the number of brains and regions analyzed so far is not sufficient to fully investigate the whole magnitude of changes in gene expression and exon usage or the full range of eQTLs. Although the application of RNA sequencing technology allowed more in-depth analysis of the transcriptome, including some low-expressing transcripts and their associated spatiotemporal patterns, it also carried some problems. The hardest one to circumvent was related to the differences at the genome annotation level. Since the genome of these species have are not annotated with a similar level of detail, it is difficult to accurately calculate differential expression among them. Finally, while specific patterns of expression are often linked to equally specialized

biological processes, it is important to notice that the relationship between mRNA and protein levels is not always linear or translated into apparent phenotypic differences. As these concerns are addressed in the future, it will be possible to uncover additional insights into the transcriptional foundations of human brain development and evolution.

The findings presented in this thesis open the door to several future research directions. Since the mRNA levels during development are so dynamically regulated, it will be interesting to study the expression of noncoding RNAs and try to correlate their expression and putative regulation of some of the co-expression networks that were found. It will be also interesting to analyze if the human-specific patterns can be explained by differences in miRNA or lncRNA expression. The extending of the interspecies comparison to developmental stages will be of the utmost relevance because it will likely reveal several human-specific patterns of expression that shape human-specific features of brain development.

CHAPTER IV

REFERENCES

Abrahams, B., D. Tentler, J. Perederiy, M. Oldham, G. Coppola and D. Geschwind (2007). "Genome-wide analyses of human perisylvian cerebral cortical patterning." Proceedings of the National Academy of Sciences of the United States of America **104**(45): 17849-17854.

Adler, D., Murdoch, D. (2010). “rgl: 3D visualization device system (OpenGL)”. R package, version 0.92.794. http://CRAN.R-project.org/package=rgl

Allen, N. C., S. Bagade, M. B. McQueen, J. P. A. Ioannidis, F. K. Kavvoura, M. J. Khoury, R. E. Tanzi and L. Bertram (2008). "Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database." Nature Genetics **40**(7): 827-834.

Amunts, K., A. Schleicher, A. Ditterich and K. Zilles (2003). "Broca's region: cytoarchitectonic asymmetry and developmental changes." Journal of Comparative Neurology **465**(1): 72-89.

Anders S, Huber W. (2010). “Differential expression analysis for sequence count data”. Genome Biology **11**(10):R106.

Andersen, S. L. (2003). "Trajectories of brain development: point of vulnerability or window of opportunity?" Neuroscience and Biobehavioral Reviews **27**(1-2): 3-18.

Anderson, J. M., R. Gilmore, S. Roper, B. Crosson, R. M. Bauer, S. Nadeau, D. Q. Beversdorf, J. Cibula, M. Rogish, 3rd, S. Kortencamp, J. D. Hughes, L. J. Gonzalez Rothi and K. M. Heilman (1999). "Conduction aphasia and the arcuate fasciculus: A reexamination of the Wernicke-Geschwind model." Brain Lang **70**(1): 1-12.

Angevine, J. B., Jr. and R. L. Sidman (1961). "Autoradiographic study of cell migration during histogenesis of cerebral cortex in the mouse." Nature **192**: 766-768.

Ariani, F., G. Hayek, D. Rondinella, R. Artuso, M. Mencarelli, A. Spanhol-Rosseto, M. Pollazzon, S. Buoni, O. Spiga, S. Ricciardi, I. Meloni, I. Longo, F. Mari, V. Broccoli, M. Zappella and A. Renieri (2008). "FOXG1 is responsible for the congenital variant of Rett syndrome." The American Journal of Human Genetics **83**(1): 89-93.

Babbitt, C. C., O. Fedrigo, A. D. Pfefferle, A. P. Boyle, J. E. Horvath, T. S. Furey and G. A. Wray (2010). "Both Noncoding and Protein-Coding RNAs Contribute to Gene Expression Evolution in the Primate Brain." Genome Biology and Evolution **2**: 67-79.

Barton, A. J., R. C. Pearson, A. Najlerahim and P. J. Harrison (1993). "Pre- and postmortem influences on brain RNA." Journal of Neurochemistry **61**(1): 1-11.

Bayer, S. A. and J. Altman (1987). "Directions in Neurogenetic Gradients and Patterns of Anatomical Connections in the Telencephalon." Progress in Neurobiology **29**(1): 57-106.

Bejerano, G., D. Haussler and M. Blanchette (2004). "Into the heart of darkness: large-scale clustering of human non-coding DNA." Bioinformatics **20**: 40-48.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." Journal of the Royal Statistical Society. Series B (Methodological) **57**(1): 289-300.

Benovoy, D., T. Kwan and J. Majewski (2008). "Effect of polymorphisms within probe-target sequences on olignonucleotide microarray experiments." Nucleic Acids Research **36**(13): 4417-4423.

Bianchi, S., C. D. Stimpson, A. L. Bauernfeind, S. J. Schapiro, W. B. Baze, M. J. McArthur, E. Bronson, W. D. Hopkins, K. Semendeferi, B. Jacobs, P. R. Hof and C. C. Sherwood (2012). "Dendritic Morphology of Pyramidal Neurons in the Chimpanzee Neocortex: Regional Specializations and Comparison to Humans." Cerebral Cortex.

Blekhman, R., A. Oshlack, A. E. Chabot, G. K. Smyth and Y. Gilad (2008). "Gene Regulation in Primates Evolves under Tissue-Specific Selection Pressures." Plos Genetics **4**(11): e1000271

Blencowe, B. J. (2006). "Alternative splicing: New insights from global analyses." Cell **126**(1): 37-47.

Bogin, B. (1988). “Patterns of human growth.” Cambridge, England; New York: Cambridge University Press.

Braak, H., E. Braak and J. Bohl (1993). "Staging of Alzheimer-Related Cortical Destruction." European Neurology **33**(6): 403-408.

Preuss, T. M. (1995). "Do Rats Have Prefrontal Cortex - the Rose-Woolsey-Akert Program Reconsidered." Journal of Cognitive Neuroscience **7**(1): 1-24.

Britanova, O., C. de Juan Romero, A. Cheung, K. Y. Kwan, M. Schwark, A. Gyorgy, T. Vogel, S. Akopov, M. Mitkovski, D. Agoston, N. Sestan, Z. Molnar and V. Tarabykin (2008). "Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex." Neuron **57**(3): 378-392.

Broca, P. (1861). "Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche." Bulletin de la Société d'Anthropologie **2**: 235–238.

Brodmann, K. (1909). “Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues”. Leipzig: Barth Verlag.

Brodmann, K. (1909). “Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues”. Leipzig: Barth Verlag.

Buxhoeveden, D. P., A. E. Switala, E. Roy, M. Litaker and M. F. Casanova (2001). "Morphological differences between minicolumns in human and nonhuman primate cortex." American Journal of Physical Anthropology **115**(4): 361-371.

Caceres, M., J. Lachuer, M. A. Zapala, J. C. Redmond, L. Kudo, D. H. Geschwind, D. J. Lockhart, T. M. Preuss and C. Barlow (2003). "Elevated gene expression levels distinguish human from non-human primate brains." Proceedings of the National Academy of Sciences of the United States of America **100**(22): 13030-13035.

Campbell, D. B., J. S. Sutcliffe, P. J. Ebert, R. Militerni, C. Bravaccio, S. Trillo, M. Elia, C. Schneider, R. Melmed, R. Sacco, A. M. Persico and P. Levitt (2006). "A genetic variant that disrupts MET transcription is associated with autism." Proceedings of the National Academy of Sciences of the United States of America **103**(45): 16834-16839.

Carroll, S. B. (2003). "Genetics and the making of Homo sapiens." Nature **422**(6934): 849-857.

Catani, M., F. Dell'acqua, A. Bizzi, S. J. Forkel, S. C. Williams, A. Simmons, D. G. Murphy and M. Thiebaut de Schotten (2012). "Beyond cortical localization in clinico-anatomical correlation." Cortex **48**(10): 1262-1287.

Catts, V. S., S. V. Catts, H. R. Fernandez, J. M. Taylor, E. J. Coulson and L. H. Lutze-Mann (2005). "A microarray study of post-mortem mRNA degradation in mouse brain tissue." Molecular Brain Research **138**(2): 164-177.

Chandana, S. R., M. E. Behen, C. Juhasz, O. Muzik, R. D. Rothermel, T. J. Mangner, P. K. Chakraborty, H. T. Chugani and D. C. Chugani (2005). "Significance of abnormalities in developmental trajectory and asymmetry of cortical serotonin synthesis in autism." Int J Developmental Neuroscience **23**(2-3): 171-182.

Charvet, C. J., G. F. Striedter and B. L. Finlay (2011). "Evo-devo and brain scaling: candidate developmental mechanisms for variation and constancy in vertebrate brain evolution." Brain Behavior and Evolution **78**(3): 248-257.

Chen, B., L. Schaevitz and S. McConnell (2005). "Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex." Proceedings of the National Academy of Sciences of the United States of America **102**(47): 17184-17189.

Chen, C.-H., E. D. Gutierrez, W. Thompson, M. S. Panizzon, T. L. Jernigan, L. T. Eyler, C. Fennema-Notestine, A. J. Jak, M. C. Neale, C. E. Franz, M. J. Lyons, M. D. Grant, B. Fischl, L. J. Seidman, M. T. Tsuang, W. S. Kremen and A. M. Dale (2012). "Hierarchical Genetic Organization of Human Cortical Surface Area." Science **335**(6076): 1634-1636.

Chen, D. Y., S. A. Stern, A. Garcia-Osta, B. Saunier-Rebori, G. Pollonini, D. Bambah-Mukku, R. D. Blitzer and C. M. Alberini (2011). "A critical role for IGF-II in memory consolidation and enhancement." Nature **469**(7331): 491-497.

Chen, J., M. Rasin, K. Kwan and N. Sestan (2005). "Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex." Proceedings of the National Academy of Sciences of the United States of America **102**(49): 17792-17797.

Chi, J. G., E. C. Dooling and F. H. Gilles (1977). "Left-right asymmetries of the temporal speech areas of the human fetus." Archives of Neurology **34**(6): 346.

Chikaraishi, D. M. (1979). "Complexity of cytoplasmic polyadenylated and nonpolyadenylated rat brain ribonucleic acids." Biochemistry **18**(15): 3249-3256.

Chodroff, R. A., L. Goodstadt, T. M. Sirey, P. L. Oliver, K. E. Davies, E. D. Green, Z. Molnar and C. P. Ponting (2010). "Long noncoding RNA genes: conservation of sequence and brain expression among diverse amniotes." Genome Biology **11**(7): R72.

Clancy, B., B. Kersh, J. Hyde, R. B. Darlington, K. J. S. Anand and B. L. Finlay (2007). "Web-based method for translating neurodevelopment from laboratory species to humans." Neuroinformatics **5**(1): 79-94.

Cohen, B. D., C. D. Noblin, A. J. Silverman and S. B. Penick (1968). "Functional asymmetry of the human brain." Science **162**(3852): 475-477.

Colantuoni, C., B. K. Lipska, T. Z. Ye, T. M. Hyde, R. Tao, J. T. Leek, E. A. Colantuoni, A. G. Elkahloun, M. M. Herman, D. R. Weinberger and J. E. Kleinman (2011). "Temporal dynamics and genetic control of transcription in the human prefrontal cortex." Nature **478**(7370): 519- 523.

Colantuoni, C., A. E. Purcell, C. M. L. Bouton and J. Pevsner (2000). "High throughput analysis of gene expression in the human brain." Journal of Neuroscience Research **59**(1): 1-10.

Consortium, E. P., E. Birney, J. A. Stamatoyannopoulos, A. Dutta, R. Guigo, T. R. Gingeras, E. H. Margulies, Z. Weng, M. Snyder, E. T. Dermitzakis, R. E. Thurman, M. S. Kuehn, C. M. Taylor, S. Neph, C. M. Koch, S. Asthana, A. Malhotra, I. Adzhubei, J. A. Greenbaum, R. M. Andrews, P. Flicek, P. J. Boyle, H. Cao, N. P. Carter, G. K. Clelland, S. Davis, N. Day, P. Dhami, S. C. Dillon, M. O. Dorschner, H. Fiegler, P. G. Giresi, J. Goldy, M. Hawrylycz, A. Haydock, R. Humbert, K. D. James, B. E. Johnson, E. M.

Johnson, T. T. Frum, E. R. Rosenzweig, N. Karnani, K. Lee, G. C. Lefebvre, P. A. Navas, F. Neri, S. C. Parker, P. J. Sabo, R. Sandstrom, A. Shafer, D. Vetrie, M. Weaver, S. Wilcox, M. Yu, F. S. Collins, J. Dekker, J. D. Lieb, T. D. Tullius, G. E. Crawford, S. Sunyaev, W. S. Noble, I. Dunham, F. Denoeud, A. Reymond, P. Kapranov, J. Rozowsky, D. Zheng, R. Castelo, A. Frankish, J. Harrow, S. Ghosh, A. Sandelin, I. L. Hofacker, R. Baertsch, D. Keefe, S. Dike, J. Cheng, H. A. Hirsch, E. A. Sekinger, J. Lagarde, J. F. Abril, A. Shahab, C. Flamm, C. Fried, J. Hackermuller, J. Hertel, M. Lindemeyer, K. Missal, A. Tanzer, S. Washietl, J. Korbel, O. Emanuelsson, J. S. Pedersen, N. Holroyd, R. Taylor, D. Swarbreck, N. Matthews, M. C. Dickson, D. J. Thomas, M. T. Weirauch, J. Gilbert, J. Drenkow, I. Bell, X. Zhao, K. G. Srinivasan, W. K. Sung, H. S. Ooi, K. P. Chiu, S. Foissac, T. Alioto, M. Brent, L. Pachter, M. L. Tress, A. Valencia, S. W. Choo, C. Y. Choo, C. Ucla, C. Manzano, C. Wyss, E. Cheung, T. G. Clark, J. B. Brown, M. Ganesh, S. Patel, H. Tammana, J. Chrast, C. N. Henrichsen, C. Kai, J. Kawai, U. Nagalakshmi, J. Wu, Z. Lian, J. Lian, P. Newburger, X. Zhang, P. Bickel, J. S. Mattick, P. Carninci, Y. Hayashizaki, S. Weissman, T. Hubbard, R. M. Myers, J. Rogers, P. F. Stadler, T. M. Lowe, C. L. Wei, Y. Ruan, K. Struhl, M. Gerstein, S. E. Antonarakis, Y. Fu, E. D. Green, U. Karaoz, A. Siepel, J. Taylor, L. A. Liefer, K. A. Wetterstrand, P. J. Good, E. A. Feingold, M. S. Guyer, G. M. Cooper, G. Asimenos, C. N. Dewey, M. Hou, S. Nikolaev, J. I. Montoya-Burgos, A. Loytynoja, S. Whelan, F. Pardi, T. Massingham, H. Huang, N. R. Zhang, I. Holmes, J. C. Mullikin, A. Ureta-Vidal, B. Paten, M. Seringhaus, D. Church, K. Rosenbloom, W. J. Kent, E. A. Stone, N. C. S. Program, C. Baylor College of Medicine Human Genome Sequencing, C. Washington University Genome Sequencing, I. Broad, I. Children's Hospital Oakland Research, S. Batzoglou, N. Goldman, R. C. Hardison, D. Haussler, W. Miller, A. Sidow, N. D. Trinklein, Z. D. Zhang, L. Barrera, R. Stuart, D. C. King, A. Ameur, S. Enroth, M. C. Bieda, J. Kim, A. A. Bhinge, N. Jiang, J. Liu, F. Yao, V. B. Vega, C. W. Lee, P. Ng, A. Shahab, A. Yang, Z. Moqtaderi, Z. Zhu, X. Xu, S. Squazzo, M. J. Oberley, D. Inman, M. A. Singer, T. A. Richmond, K. J. Munn, A. Rada-Iglesias, O. Wallerman, J. Komorowski, J. C. Fowler, P. Couttet, A. W. Bruce, O. M. Dovey, P. D. Ellis, C. F. Langford, D. A. Nix, G. Euskirchen, S. Hartman, A. E. Urban, P. Kraus, S. Van Calcar, N. Heintzman, T. H. Kim, K. Wang, C. Qu, G. Hon, R. Luna, C. K. Glass, M. G. Rosenfeld, S. F. Aldred, S. J. Cooper, A. Halees, J. M. Lin, H. P. Shulha, X. Zhang, M. Xu, J. N. Haidar, Y. Yu, Y. Ruan, V. R. Iyer, R. D. Green, C. Wadelius, P. J. Farnham, B. Ren, R. A. Harte, A. S. Hinrichs, H. Trumbower, H. Clawson, J. Hillman-Jackson, A. S. Zweig, K. Smith, A. Thakkapallayil, G. Barber, R. M. Kuhn, D. Karolchik, L. Armengol, C. P. Bird, P. I. de Bakker, A. D. Kern, N. Lopez-Bigas, J. D. Martin, B. E. Stranger, A. Woodroffe, E. Davydov, A. Dimas, E. Eyras, I. B. Hallgrimsdottir, J. Huppert, M. C. Zody, G. R. Abecasis, X. Estivill, G. G. Bouffard, X. Guan, N. F. Hansen, J. R. Idol, V. V.

Maduro, B. Maskeri, J. C. McDowell, M. Park, P. J. Thomas, A. C. Young, R. W. Blakesley, D. M. Muzny, E. Sodergren, D. A. Wheeler, K. C. Worley, H. Jiang, G. M. Weinstock, R. A. Gibbs, T. Graves, R. Fulton, E. R. Mardis, R. K. Wilson, M. Clamp, J. Cuff, S. Gnerre, D. B. Jaffe, J. L. Chang, K. Lindblad-Toh, E. S. Lander, M. Koriabine, M. Nefedov, K. Osoegawa, Y. Yoshinaga, B. Zhu and P. J. de Jong (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project." Nature **447**(7146): 799-816.

Crow, T. J. (2008). "The 'big bang' theory of the origin of psychosis and the faculty of language." Schizophrenia Research **102**(1-3): 31-52.

Cullen, T. J., M. A. Walker, S. L. Eastwood, M. M. Esiri, P. J. Harrison and T. J. Crow (2006). "Anomalies of asymmetry of pyramidal cell density and structure in dorsolateral prefrontal cortex in schizophrenia." British Journal of Psychiatry **188**: 26-31.

Dax, M. (1865). "Lésions de la moitié gauche de l'encéphale coincidant avec l'oubli des signes de la pensée (lu au Congrés méridional tenu à Montpellier en 1836)." Gaz Heb Méd Chi.

de la Grange, P., L. Gratadou, M. Delord, M. Dutertre and D. Auboeuf (2010). "Splicing factor and exon profiling across human tissues." Nucleic Acids Research **38**(9): 2825-2838.

de Lucchi, R., M., B. J. Dennis and W. R. Adey (1965). Stereotaxic Atlas of the Chimpanzee Brain (Pan satyrus). University of California Press.

Deaner, R. O., K. Isler, J. Burkart and C. van Schaik (2007). "Overall Brain Size, and Not Encephalization Quotient, Best Predicts Cognitive Ability across Non-Human Primates." Brain, Behavior and Evolution **70**(2): 115-124.

DeFelipe, J. (2011). "The evolution of the brain, the human nature of cortical circuits and intellectual creativity." Frontiers in Neuroanatomy **5:** 29.

Dehay, C. and H. Kennedy (2009). "Transcriptional regulation and alternative splicing make for better brains." Neuron **62**(4): 455-457.

Elston, G. N., R. Benavides-Piccione, A. Elston, B. Zietsch, J. Defelipe, P. Manger, V. Casagrande and J. H. Kaas (2006). "Specializations of the granular prefrontal cortex of primates: implications for cognitive processing." The anatomical record. Part A, Discoveries in molecular, cellular, and evolutionary biology **288**(1): 26-35.

Enard, W., P. Khaitovich, J. Klose, S. Zollner, F. Heissig, P. Giavalisco, K. Nieselt-Struwe, E. Muchmore, A. Varki, R. Ravid, G. M. Doxiadis, R. E. Bontrop and S. Paabo (2002). "Intra- and interspecific variation in primate gene expression patterns." Science **296**(5566): 340-343.

Fertuzinhos, S., Z. Krsnik, Y. Kawasawa, M. Rasin, K. Kwan, J. Chen, M. Judas, M. Hayashi and N. Sestan (2009). "Selective depletion of molecularly defined cortical interneurons in human holoprosencephaly with severe striatal hypoplasia." Cerebral Cortex **19**(9): 2196-2207.

Finlay, B. L. and R. B. Darlington (1995). "Linked regularities in the development and evolution of mammalian brains." Science **268**(5217): 1578-1584.

Flechsig Of Leipsic, P. (1901). "Developmental (myelogenetic) localization of the cerebral cortex in the human subject." The Lancet **158**(4077): 1027-1030.

Galuske, R. A., W. Schlote, H. Bratzke and W. Singer (2000). "Interhemispheric asymmetries of the modular structure in human temporal cortex." Science **289**(5486): 1946-1949.

Gazzaniga, M. S. (2008). “Human : the science behind what makes us unique”. New York: Ecco.

Gazzaniga, M. S., R. W. Sperry and J. E. Bogen (1962). "Some Functional Effects of Sectioning Cerebral Commissures in Man." Proceedings of the National Academy of Sciences of the United States of America **48**(10): 1765-1769.

Geschwind, N. and A. M. Galaburda (1985). "Cerebral Lateralization - Biological Mechanisms, Associations, and Pathology .3. A Hypothesis and a Program for Research." Archives of Neurology **42**(7): 634-654.

Geschwind, N. and W. Levitsky (1968). "Human Brain - Left-Right Asymmetries in Temporal Speech Region." Science **161**(3837): 186-187.

Giedd, J. N., J. Blumenthal, N. O. Jeffries, F. X. Castellanos, H. Liu, A. Zijdenbos, T. Paus, A. C. Evans and J. L. Rapoport (1999). "Brain development during childhood and adolescence: a longitudinal MRI study." Nature Neuroscience **2**(10): 861-863.

Goldman-Rakic, P. S. (1998). "The cortical dopamine system: role in memory and cognition." Adv Pharmacol **42**: 707-711.

Gould, S. J. (1977). “Ontogeny and Phylogeny”. Cambridge: Belknap Press of Harvard University Press.

Grefkes, C. and G. R. Fink (2005). "The functional organization of the intraparietal sulcus in humans and monkeys." Journal of Anatomy **207**(1): 3-17.

Gregg, J. L., K. E. Brown, E. M. Mintz, H. Piontkivska and G. C. Fraizer (2010). "Analysis of gene expression in prostate cancer epithelial and interstitial stromal cells using laser capture microdissection." BMC Cancer **10:** 165.

Gu, J. Y. and X. Gu (2003). "Induced gene expression in human brain after the split from chimpanzee." Trends in Genetics **19**(2): 63-65.

Habegger, L., A. Sboner, T. A. Gianoulis, J. Rozowsky, A. Agarwal, M. Snyder and M. Gerstein (2011). "RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries." Bioinformatics **27**(2): 281-283.

Hamamori, Y., H. Y. Wu, V. Sartorelli and L. Kedes (1997). "The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist." Molecular and Cellular Biology **17**(11): 6563-6573.

Han, W., K. Y. Kwan, S. Shim, M. M. Lam, Y. Shin, X. Xu, Y. Zhu, M. Li and N. Sestan (2011). "TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract." Proceedings of the National Academy of Sciences of the United States of America **108**(7): 3041-3046.

Hansen, K. D., R. A. Irizarry and Z. J. Wu (2012). "Removing technical variability in RNA-seq data using conditional quantile normalization." Biostatistics **13**(2): 204-216.

Harder, J. A., H. F. Baker and R. M. Ridley (1998). "The role of the central cholinergic projections in cognition: implications of the effects of scopolamine on discrimination learning by monkeys." Brain Research Bulletin **45**(3): 319-326.

Harrison-Uy, S. J. and S. J. Pleasure (2012). "Wnt signaling and forebrain development." Cold Spring Harbor Perspectives in Biology **4**(7): a008094.

Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Chrast J, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigó R, Hubbard TJ. (2012). “GENCODE: the reference human genome annotation for The ENCODE Project”. Genome Research **22**(9):1760-1774.

Hawrylycz, M. J., E. S. Lein, A. L. Guillozet-Bongaarts, E. H. Shen, L. Ng, J. A. Miller, L. N. van de Lagemaat, K. A. Smith, A. Ebbert, Z. L. Riley, C. Abajian, C. F. Beckmann, A. Bernard, D. Bertagnolli, A. F. Boe, P. M. Cartagena, M. M. Chakravarty, M. Chapin, J. Chong, R. A. Dalley, B. D. Daly, C. Dang, S. Datta, N. Dee, T. A. Dolbeare, V. Faber, D. Feng, D. R. Fowler, J. Goldy, B. W. Gregor, Z. Haradon, D. R. Haynor, J. G. Hohmann, S. Horvath, R. E. Howard, A. Jeromin, J. M. Jochim, M. Kinnunen, C. Lau, E. T. Lazarz, C. Lee, T. A. Lemon, L. Li, Y. Li, J. A. Morris, C. C. Overly, P. D. Parker, S. E. Parry, M. Reding, J. J. Royall, J. Schulkin, P. A. Sequeira, C. R. Slaughterbeck, S. C. Smith, A. J. Sodt, S. M. Sunkin, B. E. Swanson, M. P. Vawter, D. Williams, P. Wohnoutka, H. R. Zielke, D. H. Geschwind, P. R. Hof, S. M. Smith, C. Koch, S. G. N. Grant and A. R. Jones (2012). "An anatomically comprehensive atlas of the adult human brain transcriptome." Nature **489**(7416): 391-399.

Hayes, T. L. and D. A. Lewis (1993). "Hemispheric differences in layer III pyramidal neurons of the anterior language area." Archives of Neurology **50**(5): 501-505.

Haygood, R., O. Fedrigo, B. Hanson, K. D. Yokoyama and G. Awray (2007). "Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution." Nature Genetics **39**(9): 1140-1144.

Heinrich, M., K. Matt, S. Lutz-Bonengel and U. Schmidt (2007). "Successful RNA extraction from various human postmortem tissues." International Journal of Legal Medicine **121**(2): 136-142.

Heinzen, E. L., D. L. Ge, K. D. Cronin, J. M. Maia, K. V. Shianna, W. N. Gabriel, K. A. Welsh-Bohmer, C. M. Hulette, T. N. Denny and D. B. Goldstein (2008). "Tissue-Specific Genetic Control of Splicing: Implications for the Study of Complex Traits." Plos Biology **6**(12): 2869-2879.

Hekstra, D., A. R. Taussig, M. Magnasco and F. Naef (2003). "Absolute mRNA concentrations from sequence-specific calibration of oligonucleotide arrays." Nucleic Acids Research **31**(7): 1962-1968.

Herculano-Houzel, S. (2011). "Brains matter, bodies maybe not: the case for examining neuron numbers irrespective of body size." New Perspectives on Neurobehavioral Evolution **1225**: 191-199.

Hevner, R., L. Shi, N. Justice, Y. Hsueh, M. Sheng, S. Smiga, A. Bulfone, A. Goffinet, A. Campagnoni and J. Rubenstein (2001). "Tbr1 regulates differentiation of the preplate and layer 6." Neuron **29**(2): 353-366.

Hill, J., T. Inder, J. Neil, D. Dierker, J. Harwell and D. Van Essen (2010). "Similar patterns of cortical expansion during human development and evolution." Proceedings of the National Academy of Sciences of the United States of America **107**(29): 13135-13140.

Hill, R. S. and C. A. Walsh (2005). "Molecular insights into human brain evolution." Nature **437**(7055): 64-67.

Hindson BJ, Ness, KD, Masquelier, DA, Belgrader, P, Heredia, NJ, Makarewicz, AJ, Bright, IJ, Lucero, MY, Hiddesen, AL, Legler, TC, Kitano, TK, Hodel, MR, Petersen, JF, Wyatt, PW, Steenblock, ER, Shah, PH, Bousse, LJ, Troup, CB, Mellen, JC, Wittman, DK, Erndt, NG, Cauley, TH, Koehler, RT, So, AP, Dube, S, Rose, KA, Montesclaros, L, Wang, S, Stumbo, DP, Hodges, SP, Romine, S, Milanovich, FP, White, HE, Regan, JF, Karlin-

Neuman, GA, Hindson, CM, Saxonov, S, Colston, BW. (2011). High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Analytical Chemistry **83**(22): 8604–8610.

Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, Hillman-Jackson J, Kuhn RM, Pedersen JS, Pohl A, Raney BJ, Rosenbloom KR, Siepel A, Smith KE, Sugnet CW, Sultan-Qurraie A, Thomas DJ, Trumbower H, Weber RJ, Weirauch M, Zweig AS, Haussler D, Kent WJ. (2006). "The UCSC Genome Browser Database: update 2006". Nucleic Acids Res **34**: D590–D598.

His, W. (1904). "Die Entwickelung des menschlichen Gehirns während der ersten Monate. Untersuchungsergebnisse". Leipzig: Hirzel Verlag.

Howard, B. M., Z. Mo, R. Filipovic, A. R. Moore, S. D. Antic and N. Zecevic (2008). "Radial glia cells in the developing human brain." Neuroscientist **14**(5): 459-473.

Hu, Z. J., J. H. Hung, Y. Wang, Y. C. Chang, C. L. Huang, M. Huyck and C. DeLisi (2009). "VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology." Nucleic Acids Research **37**: W115-W121.

Huang, D. W., B. T. Sherman and R. A. Lempicki (2009). "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists." Nucleic Acids Research **37**(1): 1-13.

Huang, D. W., B. T. Sherman and R. A. Lempicki (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nature Protocols **4**(1): 44-57.

Huffaker, S. J., J. S. Chen, K. K. Nicodemus, F. Sambataro, F. Yang, V. Mattay, B. K. Lipska, T. M. Hyde, J. Song, D. Rujescu, I. Giegling, K. Mayilyan, M. J. Proust, A. Soghoyan, G. Caforio, J. H. Callicott, A. Bertolino, A. Meyer-Lindenberg, J. Chang, Y. J. Ji, M. F. Egan, T. E. Goldberg, J. E. Kleinman, B. Lu and D. R. Weinberger (2009). "A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia." Nature Medicine **15**(5): 509-518.

Huttenlocher, P. R. and A. S. Dabholkar (1997). "Regional differences in synaptogenesis in human cerebral cortex." Journal of Comparative Neurology **387**(2): 167-178.

Insel, T. (2010). "Rethinking schizophrenia." Nature **468**(7321): 187-193.

Ip, B., I. Wappler, H. Peters, S. Lindsay, G. Clowry and N. Bayatti (2010). "Investigating gradients of gene expression involved in early human cortical development." Journal of Anatomy **217**(4): 300-311.

Ip, B. K., I. Wappler, H. Peters, S. Lindsay, G. J. Clowry and N. Bayatti (2010). "Investigating gradients of gene expression involved in early human cortical development." Journal of Anatomy **217**(4): 300-311.

Irizarry, R., S. Ooi, Z. Wu and J. Boeke (2003). "Use of mixture models in a microarray-based screening procedure for detecting differentially represented yeast mutants." Statistical Applications in Genetics and Molecular Biology **2**: Article1.

Jamain, S., H. Quach, C. Betancur, M. Råstam, C. Colineaux, I. C. Gillberg, H. Soderstrom, B. Giros, M. Leboyer, C. Gillberg, T. Bourgeron and P. A. R. I. S. Study (2003). "Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism." Nature Genetics **34**(1): 27-29.

Jerison, H. J. (1973). “Evolution of the brain and intelligence”. New York, Academic Press.

Johnson, M. B., Y. I. Kawasawa, C. E. Mason, Z. Krsnik, G. Coppola, D. Bogdanovic, D. H. Geschwind, S. M. Mane, M. W. State and N. Sestan (2009). "Functional and evolutionary insights into human brain development through global transcriptome analysis." Neuron **62**(4): 494-509.

Johnson, M. H. (2001). "Functional brain development in humans." Nature Reviews Neuroscience **2**(7): 475-483.

Johnson, WE, Rabinovic, A, and Li, C (2007). Adjusting batch effects in microarray expression data using Empirical Bayes methods. Biostatistics **8**(1):118-127.

Jolles, D. D. and E. A. Crone (2012). "Training the developing brain: a neurocognitive perspective." Frontiers in Human Neuroscience **6:** 76.

Joutel, A., C. Corpechot, A. Ducros, K. Vahedi, H. Chabriat, P. Mouton, S. Alamowitch, V. Domenga, M. Cecillion, E. Marechal, J. Maciazek, C. Vayssiere, C. Cruaud, E. A. Cabanis, M. M. Ruchoux, J. Weissenbach, J. F. Bach, M. G. Bousser and E. TournierLasserve (1996). "Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia." Nature **383**(6602): 707-710.

Judas, M. (2010). “Prenatal Development of the Human Fetal Telencephalon. Fetal MRI”. Heidelberg: Springer Verlag.

Kaas, J. H. (2012). "The evolution of neocortex in primates." Evolution of the Primate Brain: From Neuron to Behavior **195**: 91-102.

Kaas, J. H. and T. M. Preuss (2007). “Evolution of Nervous Sytems”. Academic Press

Kalinka, A. T., K. M. Varga, D. T. Gerrard, S. Preibisch, D. L. Corcoran, J. Jarrells, U. Ohler, C. M. Bergman and P. Tomancak (2010). "Gene expression divergence recapitulates the developmental hourglass model." Nature **468**(7325): 811-U102.

Kang, H. J., Y. I. Kawasawa, F. Cheng, Y. Zhu, X. Xu, M. Li, A. M. Sousa, M. Pletikos, K. A. Meyer, G. Sedmak, T. Guennel, Y. Shin, M. B. Johnson, Z. Krsnik, S. Mayer, S. Fertuzinhos, S. Umlauf, S. N. Lisgo, A. Vortmeyer, D. R. Weinberger, S. Mane, T. M. Hyde, A. Huttner, M. Reimers, J. E. Kleinman and N. Sestan (2011). "Spatio-temporal transcriptome of the human brain." Nature **478**(7370): 483-489.

Kasprian, G., G. Langs, P. C. Brugger, M. Bittner, M. Weber, M. Arantes and D. Prayer (2011). "The prenatal origin of hemispheric asymmetry: an in utero neuroimaging study." Cerebral Cortex **21**(5): 1076-1083.

Kennedy, H. and C. Dehay (2012). "Self-organization and interareal networks in the primate cortex." Progress in Brain Research **195**: 341-360.

Kent WJ. (2002). “BLAT--the BLAST-like alignment tool”. Genome Research: **12**(4):656-64.

Khaitovich, P., B. Muetzel, X. W. She, M. Lachmann, I. Hellmann, J. Dietzsch, S. Steigele, H. H. Do, G. Weiss, W. Enard, F. Heissig, T. Arendt, K. Nieselt-Struwe, E. E. Eichler and S. Paabo (2004). "Regional patterns of gene expression in human and chimpanzee brains." Genome Research **14**(8): 1462-1473.

King, M. and A. Wilson (1975). "Evolution at two levels in humans and chimpanzees." Science **188**(4184): 107-116.

Knoth, R., I. Singec, M. Ditter, G. Pantazis, P. Capetian, R. Meyer, V. Horvat, B. Volk and G. Kempermann (2010). "Murine features of neurogenesis in the human hippocampus across the lifespan from 0 to 100 years." PLoS One **5**(1): e8809.

Kostovic, I. and M. Judas (2002). "Correlation between the sequential ingrowth of afferents and transient patterns of cortical lamination in preterm infants." Anatomical Record **267**(1): 1-6.

Kostovic, I. and M. Judas (2006). "Prolonged coexistence of transient and permanent circuitry elements in the developing cerebral cortex of fetuses and preterm infants." Developmental Medicine & Child Neurology **48**(5): 388-393.

Kostovic, I. and M. Judas (2009). “The Cognitive Neurosciences”. M. S. Gazzaniga, MIT Press.

Kostovic, I., M. Judas, Z. Petanjek and G. Simic (1995). "Ontogenesis of goal-directed behavior: anatomo-functional considerations." International Journal of Psychophysiology **19**(2): 85-102.

Kriegstein, A., S. Noctor and V. Martínez-Cerdeño (2006). "Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion." Nature Reviews Neuroscience **7**(11): 883-890.

Kumar, R. A. and S. L. Christian (2009). "Genetics of autism spectrum disorders." Current Neurology and Neuroscience Reports **9**(3): 188-197.

Kuss, A. W. and W. Chen (2008). "MicroRNAs in brain function and disease." Current Neurology and Neuroscience Reports **8**(3): 190-197.

Kwan, K., M. Lam, Z. Krsnik, Y. Kawasawa, V. Lefebvre and N. Sestan (2008). "SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons." Proceedings of the National Academy of Sciences of the United States of America **105**(41): 16021-16026.

Kwan, Kenneth Y., Mandy M. S. Lam, Matthew B. Johnson, U. Dube, S. Shim, M.-R. Rašin, André M. M. Sousa, S. Fertuzinhos, J.-G. Chen, Jon I. Arellano, Daniel W. Chan, M. Pletikos, L. Vasung, David H. Rowitch, Eric J. Huang, Michael L. Schwartz, R. Willemsen, Ben A. Oostra, P. Rakic, M. Heffer, I. Kostović, M. Judaš and N. Šestan (2012). "Species-Dependent Posttranscriptional Regulation of NOS1 by FMRP in the Developing Cerebral Cortex." Cell **149**(4): 899-911.

Lambert, N., M. A. Lambot, A. Bilheu, V. Albert, Y. Englert, F. Libert, J. C. Noel, C. Sotiriou, A. K. Holloway, K. S. Pollard, V. Detours and P. Vanderhaeghen (2011). "Genes Expressed in Specific Areas of the Human Fetal Cerebral Cortex Display Distinct Patterns of Evolution." Plos One **6**(3): e17753.

Langfelder, P., Luo, R., Oldham, M. D., and S. Horvath (2011). "Is my network module preserved and reproducible?" PLoS Computational Biology **7**: e1001057.

Langfelder, P. and S. Horvath (2008). "WGCNA: an R package for weighted correlation network analysis." BMC Bioinformatics **9**: 559.

Lawson-Yuen, A., J. S. Saldivar, S. Sommer and J. Picker (2008). "Familial deletion within NLGN4 associated with autism and Tourette syndrome." Eur J Hum Genet **16**(5): 614-618.

Lehtinen, M. K., M. W. Zappaterra, X. Chen, Y. J. Yang, A. D. Hill, M. Lun, T. Maynard, D. Gonzalez, S. Kim, P. Ye, A. J. D'Ercole, E. T. Wong, A. S. Lamantia and C. A. Walsh (2011). "The cerebrospinal fluid provides a proliferative niche for neural progenitor cells." Neuron **69**(5): 893-905.

Lewis, D. A. and P. Levitt (2002). "Schizophrenia as a disorder of neurodevelopment." Annual Review of Neuroscience **25**: 409-432.

Li, B., V. Ruotti, R. M. Stewart, J. A. Thomson and C. N. Dewey (2010). "RNA-Seq gene expression estimation with read mapping uncertainty." Bioinformatics **26**(4): 493-500.

Licatalosi, D. D. and R. B. Darnell (2006). "Splicing regulation in neurologic disease." Neuron **52**(1): 93-101.

Liu, C., L. Cheng, J. A. Badner, D. Zhang, D. W. Craig, M. Redman and E. S. Gershon (2010). "Whole-genome association mapping of gene expression in the human prefrontal cortex." Molecular Psychiatry **15**(8): 779-784.

Liu, X. L., M. Somel, L. Tang, Z. Yan, X. Jiang, S. Guo, Y. Yuan, L. He, A. Oleksiak, Y. Zhang, N. Li, Y. H. Hu, W. Chen, Z. L. Qiu, S. Paabo and P. Khaitovich (2012). "Extension of cortical synaptic development distinguishes humans from chimpanzees and macaques." Genome Research **22**(4): 611-622.

Lockhart, D. J. and C. Barlow (2001). "Expressing what's on your mind: DNA arrays and the brain." Nature Reviews Neuroscience **2**(1): 63-68.

Lui, Jan H., David V. Hansen and Arnold R. Kriegstein (2011). "Development and Evolution of the Human Neocortex." Cell **146**(1): 18-36.

MacLeod, C. E., K. Zilles, A. Schleicher, J. K. Rilling and K. R. Gibson (2003). "Expansion of the neocerebellum in Hominoidea." Journal of Human Evolution **44**(4): 401-429.

Marco-Sola, S, Sammeth, M, Guigó, R, Ribeca, P. (2012). “The GEM mapper: fast, accurate and versatile alignment by filtration”. Nature Methods **9**: 1185–1188.

McKenna, W. L., J. Betancourt, K. A. Larkin, B. Abrams, C. Guo, J. L. Rubenstein and B. Chen (2011). "Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development." Journal of Neuroscience **31**(2): 549-564.

Mega, M. S. (2000). "The cholinergic deficit in Alzheimer's disease: impact on cognition, behaviour and function." International Journal of Neuropsychopharmacology **3**(7): 3-12.

Meneses, A. and E. Hong (1995). "Effect of fluoxetine on learning and memory involves multiple 5-HT systems." Pharmacology Biochemestry and Behavior **52**(2): 341-346.

Meyer, G. (2007). "Genetic control of neuronal migrations in human cortical development." Advances in Anatomy Embryology and Cell Biology **189**. Springer

Meyer-Lindenberg, A. and D. R. Weinberger (2006). "Intermediate phenotypes and genetic mechanisms of psychiatric disorders." Nature Reviews Neuroscience **7**(10): 818-827.

Meyer-Lindenberg, A. and D. R. Weinberger (2006). "Intermediate phenotypes and genetic mechanisms of psychiatric disorders." Nature Reviews Neuroscience **7**(10): 818-827.

Mikkelsen, T. S., L. W. Hillier, E. E. Eichler, M. C. Zody, D. B. Jaffe, S. P. Yang, W. Enard, I. Hellmann, K. Lindblad-Toh, T. K. Altheide, N. Archidiacono, P. Bork, J. Butler, J. L. Chang, Z. Cheng, A. T. Chinwalla, P. deJong, K. D. Delehaunty, C. C. Fronick, L. L. Fulton, Y. Gilad, G. Glusman, S. Gnerre, T. A. Graves, T. Hayakawa, K. E. Hayden, X. Q. Huang, H. K. Ji, W. J. Kent, M. C. King, E. J. Kulbokas, M. K. Lee, G. Liu, C. Lopez-Otin, K. D. Makova, O. Man, E. R. Mardis, E. Mauceli, T. L. Miner, W. E. Nash, J. O. Nelson, S. Paabo, N. J. Patterson, C. S. Pohl, K. S. Pollard, K. Prufer, X. S. Puente, D. Reich, M. Rocchi, K. Rosenbloom, M. Ruvolo, D. J. Richter, S. F. Schaffner, A. F. A. Smit, S. M. Smith, M. Suyama, J. Taylor, D. Torrents, E. Tuzun, A. Varki, G. Velasco, M. Ventura, J. W. Wallis, M. C. Wendl, R. K. Wilson, E. S. Lander, R. H. Waterston and C. S. A. Consortium (2005). "Initial sequence of the chimpanzee genome and comparison with the human genome." Nature **437**(7055): 69-87.

Miller, D. J., T. Duka, C. D. Stimpson, S. J. Schapiro, W. B. Baze, M. J. McArthur, A. J. Fobbs, A. M. M. Sousa, N. Sestan, D. E. Wildman, L. Lipovich, C. W. Kuzawa, P. R. Hof and C. C. Sherwood (2012). "Prolonged myelination in human neocortical evolution." Proceedings of the National Academy of Sciences of the United States of America **109**(41): 16480-16485.

Mirnics, K., F. A. Middleton, A. Marquez, D. A. Lewis and P. Levitt (2000). "Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex." Neuron **28**(1): 53-67.

Molnar, Z. and G. Clowry (2012). "Cerebral cortical development in rodents and primates." Progress in Brain Research **195**: 45-70.

Mortazavi, A., B. A. Williams, K. Mccue, L. Schaeffer and B. Wold (2008). "Mapping and quantifying mammalian transcriptomes by RNA-Seq." Nature Methods **5**(7): 621-628.

Myers, A. J., J. R. Gibbs, J. A. Webster, K. Rohrer, A. Zhao, L. Marlowe, M. Kaleem, D. Leung, L. Bryden, P. Nath, V. L. Zismann, K. Joshipura, M. J. Huentelman, D. Hu-Lince, K. D. Coon, D. W. Craig, J. V. Pearson, P. Holmans, C. B. Heward, E. M. Reiman, D. Stephan and J. Hardy (2007). "A survey of genetic human cortical gene expression." Nature Genetics **39**(12): 1494-1499.

Naughton, M., J. B. Mulrooney and B. E. Leonard (2000). "A review of the role of serotonin receptors in psychiatric disorders." Human Psychopharmacology **15**(6): 397-415.

Newbury, D. and A. Monaco (2010). "Genetic advances in the study of speech and language disorders." Neuron **68**(2): 309-320.

Nieoullon, A. (2002). "Dopamine and the regulation of cognition and attention." Progress in Neurobiology **67**(1): 53-83.

Okoniewski, M. J. and C. J. Miller (2008). "Comprehensive Analysis of Affymetrix Exon Arrays Using BioConductor." Plos Computational Biology **4**(2): e6.

O'Leary, D. D. and S. Sahara (2008). "Genetic regulation of arealization of the neocortex." Current Opinion in Neurobiology **18**(1): 90-100.

O'Rahilly, R. and F. Müller (2006). “The Embryonic Human Brain”. Wiley-Liss.

Paus, T., M. Keshavan and J. Giedd (2008). "Why do many psychiatric disorders emerge during adolescence?" Nature Reviews Neuroscience **9**(12): 947-957.

Paxinos, G., G. Halliday, C. Watson, Y. Koutcherov and H. Wang (2007). “Atlas of the Developing Mouse Brain”. Academic Press-Elsevier.

Perroud, N., R. Uher, M. Ng, M. Guipponi, J. Hauser, N. Henigsberg, W. Maier, O. Mors, M. Gennarelli, M. Rietschel, D. Souery, M. Dernovsek, A. Stamp, M. Lathrop, A. Farmer, G. Breen, K. Aitchison, C. Lewis, I. Craig and P. McGuffin (2010). "Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project." The Pharmacogenomics Journal **12**(1):68-77.

Petanjek, Z., M. Judas, I. Kostović and H. Uylings (2008). "Lifespan alterations of basal dendritic trees of pyramidal neurons in the human prefrontal cortex: a layer-specific pattern." Cerebral Cortex **18**(4): 915-929.

Petanjek, Z. and I. Kostovic (2012). "Epigenetic regulation of fetal brain development and neurocognitive outcome." Proceedings of the National Academy of Sciences of the United States of America **109**(28): 11062-11063.

Petrides, M., G. Cadoret and S. Mackey (2005). "Orofacial somatomotor responses in the macaque monkey homologue of Broca's area." Nature **435**(7046): 1235-1238.

Pinto, L., D. Drechsel, M. T. Schmid, J. Ninkovic, M. Irmler, M. S. Brill, L. Restani, L. Gianfranceschi, C. Cerri, S. N. Weber, V. Tarabykin, K. Baer, F. Guillemot, J. Beckers, N. Zecevic, C. Dehay, M. Caleo, H. Schorle and M. Götz (2009). "AP2gamma regulates basal progenitor fate in a region- and layer-specific manner in the developing cortex." Nature Neuroscience **12**(10): 1229-1237.

Poliakov, G. I. (1949). “Cytoarchitectonics of the Human Cerebral Cortex”. Medgiz.

Poliakov, G. I. (1979). “Entwicklung der Neuronen der menschlichen Grosshirnrinde (Herausgegeben von B. Schönheit)”. Leipzig: Verlag Gerorg Thieme.

Pollard, K. S., S. R. Salama, B. King, A. D. Kern, T. Dreszer, S. Katzman, A. Siepel, J. S. Pedersen, G. Bejerano, R. Baertsch, K. R. Rosenbloom, J. Kent and D. Haussler (2006). "Forces shaping the fastest evolving regions in the human genome." Plos Genetics **2**(10): 1599-1611.

Ponjavic, J., P. L. Oliver, G. Lunter and C. P. Ponting (2009). "Genomic and Transcriptional Co-Localization of Protein-Coding and Long Non-Coding RNA Pairs in the Developing Brain." Plos Genetics **5**(8): e1000617.

Prabhakar, S., J. P. Noonan, S. Paabo and E. M. Rubin (2006). "Accelerated evolution of conserved noncoding sequences in humans." Science **314**(5800): 786-786.

Preuss, T. (2000). “What's human about the human brain?”. The New Cognitive Neurosciences. Second Edition. M. S. Gazzaniga. Cambridge, MIT Press.

Preuss, T., M. Cáceres, M. Oldham and D. Geschwind (2004). "Human brain evolution: insights from microarrays." Nature Reviews in Genetics **5**(11): 850-860.

Preuss, T. M. (2011). "The human brain: rewired and running hot." Annals of the New York Academy of Sciences **1225 Suppl 1**: E182-191.

Preuss, T. M. and G. Q. Coleman (2002). "Human-specific organization of primary visual cortex: alternating compartments of dense Cat-301 and calbindin immunoreactivity in layer 4A." Cerebral Cortex **12**(7): 671-691.

Preuss, T. M. and P. S. Goldman-Rakic (1991). "Architectonics of the parietal and temporal association cortex in the strepsirhine primate Galago compared to the anthropoid primate Macaca." Journal of Comparative Neurology **310**(4): 475-506.

Priddle, T. H. and T. J. Crow (2013). "The protocadherin 11X/Y (PCDH11X/Y) gene pair as determinant of cerebral asymmetry in modern Homo sapiens." Annals of the New York Academy of Sciences [Epub ahead of print].

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly and P. C. Sham (2007). "PLINK: A tool set for whole-genome association and population-based linkage analyses." American Journal of Human Genetics **81**(3): 559-575.

Qi, H. B., L. X. Xing, K. J. Zhan, X. C. Gao, Z. J. Zheng, S. P. Huang, Y. Guo and F. C. Zhang (2009). "Positive association of neuroligin-4 gene with nonspecific mental retardation in the Qinba Mountains Region of China." Psychiatric Genetics **19**(1): 1-5.

R Development Core Team (2009). “R: A language and environment for statistical computing”. Vienna, R Foundation for Statistical Computing.

Raghanti, M. A., C. D. Stimpson, J. L. Marcinkiewicz, J. M. Erwin, P. R. Hof and C. C. Sherwood (2008). "Cholinergic innervation of the frontal cortex: differences among humans, chimpanzees, and macaque monkeys." Journal of Comparative Neurology **506**(3): 409-424.

Raghanti, M. A., C. D. Stimpson, J. L. Marcinkiewicz, J. M. Erwin, P. R. Hof and C. C. Sherwood (2008). "Differences in cortical serotonergic innervation among humans, chimpanzees, and macaque monkeys: a comparative study." Cerebral Cortex **18**(3): 584-597.

Rakic, P. (1974). "Neurons in Rhesus-Monkey Visual-Cortex - Systematic Relation between Time of Origin and Eventual Disposition." Science **183**(4123): 425-427.

Rakic, P. (1988). "Specification of cerebral cortical areas." Science **241**(4862): 170-176.

Rakic, P. (2009). "Evolution of the neocortex: a perspective from developmental biology." Nature Reviews Neuroscience **10**(10): 724-735.

Rakic, P., J. P. Bourgeois, M. F. Eckenhoff, N. Zecevic and P. S. Goldman-Rakic (1986). "Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex." Science **232**(4747): 232-235.

Ramskold, D., E. T. Wang, C. B. Burge and R. Sandberg (2009). "An Abundance of Ubiquitously Expressed Genes Revealed by Tissue Transcriptome Sequence Data." Plos Computational Biology **5**(12): e1000598.

Rapoport, J. L. and N. Gogtay (2007). "Brain Neuroplasticity in Healthy, Hyperactive and Psychotic Children: Insights from Neuroimaging." Neuropsychopharmacology **33**(1): 181-197.

Rash, B. G. and E. A. Grove (2006). "Area and layer patterning in the developing cerebral cortex." Current Opinion in Neurobiology **16**(1): 25-34.

Reinius, B. and E. Jazin (2009). "Prenatal sex differences in the human brain." Mol Psychiatry **14**(11): 987-989.

Ren, C., C. H. Ren, L. Li, A. A. Goltsov and T. C. Thompson (2006). "Identification and characterization of RTVP1/GLIPR1-like genes, a novel p53 target gene cluster." Genomics **88**(2): 163-172.

Retzius, G. (1896). “Studien in der makroskopischen Morphologie”. Text und Atlas. . Stockholm, Norstedt.

Rilling, J. K. (2006). "Human and nonhuman primate brains: Are they allometrically scaled versions of the same design?" Evolutionary Anthropology **15**(2): 65-77.

Rilling, J. K., M. F. Glasser, T. M. Preuss, X. Ma, T. Zhao, X. Hu and T. E. Behrens (2008). "The evolution of the arcuate fasciculus revealed with comparative DTI." Nature Neuroscience **11**(4): 426-428.

Rilling, J. K. and T. R. Insel (1998). "Evolution of the cerebellum in primates: Differences in relative volume among monkeys, apes and humans." Brain Behavior and Evolution **52**(6): 308-314.

Roth, B. L., S. M. Hanizavareh and A. E. Blum (2004). "Serotonin receptors represent highly favorable molecular targets for cognitive enhancement in schizophrenia and other disorders." Psychopharmacology **174**(1): 17-24.

Roth, R. B., P. Hevezi, J. Lee, D. Willhite, S. M. Lechner, A. C. Foster and A. Zlotnik (2006). "Gene expression analyses reveal molecular relationships among 20 regions of the human CNS." Neurogenetics **7**(2): 67-80.

Rubenstein, J. L. (2010). "Research Review: Development of the cerebral cortex: implications for neurodevelopmental disorders." Journal of Child Psychology and Psychiatry, and allied disciplines **52**(4): 339-55.

Ryan, M. M., H. E. Lockstone, S. J. Huffaker, M. T. Wayland, M. J. Webster and S. Bahn (2006). "Gene expression analysis of bipolar disorder reveals downregulation of the ubiquitin cycle and alterations in synaptic genes." Molecular Psychiatry **11**(10): 965-978.

Sanides, F. (1969). "Comparative architectonics of the neocortex of mammals an their evolutionary interpretation." Annals of the New York Academy of Sciences **167**(1): 404-423.

Sarter, M. and V. Parikh (2005). "Choline transporters, cholinergic transmission and cognition." Nature Reviews in Neuroscience **6**(1): 48-56.

Schambra, U. (2008). “Prenatal Mouse Brain Atlas”, Springer.

Schenker, N. M., D. P. Buxhoeveden, W. L. Blackmon, K. Amunts, K. Zilles and K. Semendeferi (2008). "A comparative quantitative analysis of cytoarchitecture and minicolumnar organization in Broca's area in humans and great apes." Journal of Comparative Neurology **510**(1): 117-128.

Semendeferi, K., K. Teffer, D. P. Buxhoeveden, M. S. Park, S. Bludau, K. Amunts, K. Travis and J. Buckwalter (2011). "Spatial Organization of Neurons in the Frontal Pole Sets Humans Apart from Great Apes." Cerebral Cortex **21**(7): 1485-1497.

Semeralul, M. O., P. C. Boutros, O. Likhodi, A. B. Okey, H. H. M. Van Tol and A. H. C. Wong (2006). "Microarray analysis of the developing cortex." Journal of Neurobiology **66**(14): 1646-1658.

Sequencing, R. M. G., A. Consortium, R. A. Gibbs, J. Rogers, M. G. Katze, R. Bumgarner, G. M. Weinstock, E. R. Mardis, K. A. Remington, R. L. Strausberg, J. C. Venter, R. K. Wilson, M. A. Batzer, C. D. Bustamante, E. E. Eichler, M. W. Hahn, R. C. Hardison, K. D. Makova, W. Miller, A. Milosavljevic, R. E. Palermo, A. Siepel, J. M. Sikela, T. Attaway, S. Bell, K. E. Bernard, C. J. Buhay, M. N. Chandrabose, M. Dao, C. Davis, K. D. Delehaunty, Y. Ding, H. H. Dinh, S. Dugan-Rocha, L. A. Fulton, R. A. Gabisi, T. T. Garner, J. Godfrey, A. C. Hawes, J. Hernandez, S. Hines, M. Holder, J. Hume, S. N. Jhangiani, V. Joshi, Z. M. Khan, E. F. Kirkness, A. Cree, R. G. Fowler, S. Lee, L. R. Lewis, Z. Li, Y.-s. Liu, S. M. Moore, D. Muzny, L. V. Nazareth, D. N. Ngo, G. O. Okwuonu, G. Pai, D. Parker, H. A. Paul, C. Pfannkoch, C. S. Pohl, Y.-H. Rogers, S. J. Ruiz, A. Sabo, J. Santibanez, B. W. Schneider, S. M. Smith, E. Sodergren, A. F. Svatek, T. R. Utterback, S. Vattathil, W. Warren, C. S. White, A. T. Chinwalla, Y. Feng, A. L. Halpern, L. W. Hillier, X. Huang, P. Minx, J. O. Nelson, K. H. Pepin, X. Qin, G. G. Sutton, E. Venter, B. P. Walenz, J. W. Wallis, K. C. Worley, S.-P. Yang, S. M. Jones, M. A. Marra, M. Rocchi, J. E. Schein, R. Baertsch, L. Clarke, M. Csürös, J. Glasscock, R. A. Harris, P. Havlak, A. R. Jackson, H. Jiang, Y. Liu, D. N. Messina, Y. Shen, H. X.-Z. Song, T. Wylie, L. Zhang, E. Birney, K. Han, M. K. Konkel, J. Lee, A. F. A. Smit, B. Ullmer, H. Wang, J. Xing, R. Burhans, Z. Cheng, J. E. Karro, J. Ma, B. Raney, X. She, M. J. Cox, J. P. Demuth, L. J. Dumas, S.-G. Han, J. Hopkins, A. Karimpour-Fard, Y. H. Kim, J. R.

Pollack, T. Vinar, C. Addo-Quaye, J. Degenhardt, A. Denby, M. J. Hubisz, A. Indap, C. Kosiol, B. T. Lahn, H. A. Lawson, A. Marklein, R. Nielsen, E. J. Vallender, A. G. Clark, B. Ferguson, R. D. Hernandez, K. Hirani, H. Kehrer-Sawatzki, J. Kolb, S. Patil, L.-L. Pu, Y. Ren, D. G. Smith, D. A. Wheeler, I. Schenck, E. V. Ball, R. Chen, D. N. Cooper, B. Giardine, F. Hsu, W. J. Kent, A. Lesk, D. L. Nelson, W. E. O'Brien, K. Prüfer, P. D. Stenson, J. C. Wallace, H. Ke, X.-M. Liu, P. Wang, A. P. Xiang, F. Yang, G. P. Barber, D. Haussler, D. Karolchik, A. D. Kern, R. M. Kuhn, K. E. Smith and A. S. Zwieg (2007). "Evolutionary and Biomedical Insights from the Rhesus Macaque Genome." Science **316**(5822): 222-234.

Sherwood, C. C., F. Subiaul and T. W. Zawidzki (2008). "A natural history of the human mind: tracing evolutionary changes in brain and cognition." Journal of Anatomy **212**(4): 426-454.

Sidman, R. L. and P. Rakic (1973). "Neuronal migration, with special reference to developing human brain: a review." Brain Research **62**(1): 1-35.

Sidman, R. L. and P. Rakic (1982). “Histology and Histopathology of the Nervous System”, W. H. R. D. Adams, C.C. Thomas Publisher**:** 3-145.

Somel, M., H. Franz, Z. Yan, A. Lorenc, S. Guo, T. Giger, J. Kelso, B. Nickel, M. Dannemann, S. Bahn, M. J. Webster, C. S. Weickert, M. Lachmann, S. Paabo and P. Khaitovich (2009). "Transcriptional neoteny in the human brain." Proceedings of the National Academy of Sciences of the United States of America **106**(14): 5743-5748.

Somel, M., S. Guo, N. Fu, Z. Yan, H. Hu, Y. Xu, Y. Yuan, Z. Ning, Y. Hu, C. Menzel, H. Hu, M. Lachmann, R. Zeng, W. Chen and P. Khaitovich (2010). "MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain." Genome Research **20**(9): 1207-1218.

Somel, M., X. L. Liu, L. Tang, Z. Yan, H. Y. Hu, S. Guo, X. Jiang, X. Y. Zhang, G. H. Xu, G. C. Xie, N. Li, Y. H. Hu, W. Chen, S. Paabo and P. Khaitovich (2011). "MicroRNA-Driven Developmental Remodeling in the Brain Distinguishes Humans from Other Primates." Plos Biology **9**(12): e1001214.

Sowell, E. R., B. S. Peterson, P. M. Thompson, S. E. Welcome, A. L. Henkenius and A. W. Toga (2003). "Mapping cortical change across the human life span." Nature Neuroscience **6**(3): 309-315.

Spocter, M. A., W. D. Hopkins, S. K. Barks, S. Bianchi, A. E. Hehmeyer, S. M. Anderson, C. D. Stimpson, A. J. Fobbs, P. R. Hof and C. C. Sherwood (2012). "Neuropil distribution in the cerebral cortex differs between humans and chimpanzees." Journal of Comparative Neurology **520**(13): 2917-2929.

Srinivasan, K., D. P. Leone, R. K. Bateson, G. Dobreva, Y. Kohwi, T. Kohwi-Shigematsu, R. Grosschedl and S. K. McConnell (2012). "A network of genetic repression and derepression specifies projection fates in the developing neocortex." Proceedings of the National Academy of Sciences of the United States of America **109**(47): 19071-19078.

State, M. (2010). "The genetics of child psychiatric disorders: focus on autism and Tourette syndrome." Neuron **68**(2): 254-269.

State, M. W. (2010). "The Genetics of Child Psychiatric Disorders: Focus on Autism and Tourette Syndrome." Neuron **68**(2): 254-269.

Staudt, M. (2010). "Reorganization after pre- and perinatal brain lesions." Journal of Anatomy **217**(4): 469-474.

Steckler, T. and A. Sahgal (1995). "The role of serotonergic-cholinergic interactions in the mediation of cognitive behaviour." Behavioural Brain Research **67**(2): 165-199.

Stedman, H. H., B. W. Kozyak, A. Nelson, D. M. Thesier, L. T. Su, D. W. Low, C. R. Bridges, J. B. Shrager, N. Minugh-Purvis and M. A. Mitchell (2004). "Myosin gene mutation correlates with anatomical changes in the human lineage." Nature **428**(6981): 415-418.

Stefansson, H., R. A. Ophoff, S. Steinberg, O. A. Andreassen, S. Cichon, D. Rujescu, T. Werge, O. P. Pietiläinen, O. Mors, P. B. Mortensen, E. Sigurdsson, O. Gustafsson, M. Nyegaard, A. Tuulio-Henriksson, A. Ingason, T. Hansen, J. Suvisaari, J. Lonnqvist, T. Paunio, A. D. Børglum, A. Hartmann, A. Fink-Jensen, M. Nordentoft, D. Hougaard, B. Norgaard-Pedersen, Y. Böttcher, J. Olesen, R. Breuer, H. J. Möller, I. Giegling, H. B. Rasmussen, S. Timm, M. Mattheisen, I. Bitter, J. M. Réthelyi, B. B. Magnusdottir, T.

Sigmundsson, P. Olason, G. Masson, J. R. Gulcher, M. Haraldsson, R. Fossdal, T. E. Thorgeirsson, U. Thorsteinsdottir, M. Ruggeri, S. Tosato, B. Franke, E. Strengman, L. A. Kiemeney, I. Melle, S. Djurovic, L. Abramova, V. Kaleda, J. Sanjuan, R. de Frutos, E. Bramon, E. Vassos, G. Fraser, U. Ettinger, M. Picchioni, N. Walker, T. Toulopoulou, A. C. Need, D. Ge, J. L. Yoon, K. V. Shianna, N. B. Freimer, R. M. Cantor, R. Murray, A. Kong, V. Golimbet, A. Carracedo, C. Arango, J. Costas, E. G. Jönsson, L. Terenius, I. Agartz, H. Petursson, M. M. Nöthen, M. Rietschel, P. M. Matthews, P. Muglia, L. Peltonen, D. St Clair, D. B. Goldstein, K. Stefansson, D. A. Collier, G. R. and O. i. Psychosis (2009). "Common variants conferring risk of schizophrenia." Nature **460**(7256): 744-747.

Stranger, B. E., A. C. Nica, M. S. Forrest, A. Dimas, C. P. Bird, C. Beazley, C. E. Ingle, M. Dunning, P. Flicek, D. Koller, S. Montgomery, S. Tavare, P. Deloukas and E. T. Dermitzakis (2007). "Population genomics of human gene expression." Nature Genetics **39**(10): 1217-1224.

Südhof, T. C. (2008). "Neuroligins and neurexins link synaptic function to cognitive disease." Nature **455**(7215): 903-911.

Sun, T., C. Patoine, A. Abu-Khalil, J. Visvader, E. Sum, T. J. Cherry, S. H. Orkin, D. H. Geschwind and C. A. Walsh (2005). "Early asymmetry of gene transcription in embryonic human left and right cerebral cortex." Science **308**(5729): 1794-1798.

Sur, M. and J. L. Rubenstein (2005). "Patterning and plasticity of the cerebral cortex." Science **310**(5749): 805-810.

Sutcliffe, J. G. (1988). "mRNA in the mammalian central nervous system." Annual Review in Neuroscience **11**: 157-198.

Svenningsson, P., K. Chergui, I. Rachleff, M. Flajolet, X. Zhang, M. El Yacoubi, J. M. Vaugeois, G. G. Nomikos and P. Greengard (2006). "Alterations in 5-HT1B receptor function by p11 in depression-like states." Science **311**(5757): 77-80.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar (2011). "MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods." Molecular Biology and Evolution **28**(10): 2731-2739.

Taylor, D. C. (1969). "Differential Rates of Cerebral Maturation between Sexes and between Hemispheres - Evidence from Epilepsy." Lancet **2**(7612): 140-142.

Thatcher, R. W., R. A. Walker and S. Giudice (1987). "Human Cerebral Hemispheres Develop at Different Rates and Ages." Science **236**(4805): 1110-1113.

Toga, A. W. and P. M. Thompson (2003). "Mapping brain asymmetry." Nature Reviews Neuroscience **4**(1): 37-48.

Tripodi, M., A. Filosa, M. Armentano and M. Studer (2004). "The COUP-TF nuclear receptors regulate cell migration in the mammalian basal forebrain." Development **131**(24): 6119-6129.

Uddin, M., D. E. Wildman, G. Z. Liu, W. B. Xu, R. M. Johnson, P. R. Hof, G. Kapatos, L. I. Grossman and M. Goodman (2004). "Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles." Proceedings of the National Academy of Sciences of the United States of America **101**(9): 2957-2962.

Van Essen, D. C. (1997). "A tension-based theory of morphogenesis and compact wiring in the central nervous system." Nature **385**(6614): 313-318.

Vargha-Khadem, F., D. G. Gadian, A. Copp and M. Mishkin (2005). "FOXP2 and the neuroanatomy of speech and language." Nature Reviews Neuroscience **6**(2): 131-138.

Vawter, M., S. Evans, P. Choudary, H. Tomita, J. Meador-Woodruff, M. Molnar, J. Li, J. Lopez, R. Myers, D. Cox, S. Watson, H. Akil, E. Jones and W. Bunney (2004). "Gender-specific gene expression in post-mortem human brain: localization to sex chromosomes." Neuropsychopharmacology **29**(2): 373-384.

Visel, A., S. Prabhakar, J. A. Akiyama, M. Shoukry, K. D. Lewis, A. Holt, I. Plajzer-Frick, V. Afzal, E. M. Rubin and L. A. Pennacchio (2008). "Ultraconservation identifies a small subset of extremely constrained developmental enhancers." Nature Genetics **40**(2): 158-160.

Voineagu, I., X. C. Wang, P. Johnston, J. K. Lowe, Y. Tian, S. Horvath, J. Mill, R. M. Cantor, B. J. Blencowe and D. H. Geschwind (2011). "Transcriptomic analysis of autistic brain reveals convergent molecular pathology." Nature **474**(7351): 380-384.

Wang, J. X., K. Scully, X. Zhu, L. Cai, J. Zhang, G. G. Prefontaine, A. Krones, K. A. Ohgi, P. Zhu, I. Garcia-Bassets, F. Liu, H. Taylor, J. Lozach, F. L. Jayes, K. S. Korach, C. K. Glass, X. D. Fu and M. G. Rosenfeld (2007). "Opposing LSD1 complexes function in developmental gene activation and repression programmes." Nature **446**(7138): 882-887.

Wang, Y. Q. and B. Su (2004). "Molecular evolution of microcephalin, a gene determining human brain size." Human Molecular Genetics **13**(11): 1131-1137.

Webster, J. A., J. R. Gibbs, J. Clarke, M. Ray, W. X. Zhang, P. Holmans, K. Rohrer, A. Zhao, L. Marlowe, M. Kaleem, D. S. McCorquodale, C. Cuello, D. Leung, L. Bryden, P. Nath, V. L. Zismann, K. Joshipura, M. J. Huentelman, D. Hu-Lince, K. D. Coon, D. W. Craig, J. V. Pearson, C. B. Heward, E. M. Reiman, D. Stephan, J. Hardy, A. J. Myers and N.-N. Grp (2009). "Genetic Control of Human Brain Transcript Expression in Alzheimer Disease." American Journal of Human Genetics **84**(4): 445-458.

Weickert, C., M. Elashoff, A. Richards, D. Sinclair, S. Bahn, S. Paabo, P. Khaitovich and M. Webster (2009). "Transcriptome analysis of male-female differences in prefrontal cortical development." Molecular Psychiatry **14**(6): 558-561.

Yang, Y. H., S. Dudoit, P. Luu and T. P. Speed (2001). "Normalization for cDNA microarray data." Microarrays: Optical Technologies and Informatics **2**(23): 141-152.

Yeo, G., D. Holste, G. Kreiman and C. B. Burge (2004). "Variation in alternative splicing across human tissues." Genome Biology **5**(10): R74.

Yuan, Y., Y. P. P. Chen, S. Y. Ni, A. G. Xu, L. Tang, M. Vingron, M. Somel and P. Khaitovich (2011). "Development and application of a modified dynamic time warping algorithm (DTW-S) to analyses of primate brain expression time series." BMC Bioinformatics **12:** 347.

Zapala, M. A., I. Hovatta, J. A. Ellison, L. Wodicka, J. A. Del Rio, R. Tennant, W. Tynan, R. S. Broide, R. Helton, B. S. Stoveken, C. Winrow, D. J. Lockhart, J. F. Reilly, W. G. Young, F. E. Bloom, D. J. Lockhart and C. Barlow (2005). "Adult mouse brain gene expression patterns bear an embryologic imprint." Proceedings of the National Academy of Sciences of the United States of America **102**(29): 10357-10362.

Zhang, B. and S. Horvath (2005). "A general framework for weighted gene co-expression network analysis." Statistical Applications in Genetics and Molecular Biology **4**: Article17.

CHAPTER V

PUBLICATIONS

De acordo com o disposto no nº 2, alínea a, do artigo 31º do decreto lei nº 230/2009, utilizaram-se neste trabalho resultados já publicados no seguinte artigo:

Kang, H. J., Y. I. Kawasawa, F. Cheng, Y. Zhu, X. Xu, M. Li, A. M. M. Sousa, M. Pletikos, K. A. Meyer, G. Sedmak, T. Guennel, Y. Shin, M. B. Johnson, Z. Krsnik, S. Mayer, S. Fertuzinhos, S. Umlauf, S. N. Lisgo, A. Vortmeyer, D. R. Weinberger, S. Mane, T. M. Hyde, A. Huttner, M. Reimers, J. E. Kleinman and N. Sestan (2011). "Spatio-temporal transcriptome of the human brain." Nature **478** (7370): 483-489.

ARTICLE

doi:10.1038/nature10523

Spatio-temporal transcriptome of the human brain

Hyo Jung Kang[1]*, Yuka Imamura Kawasawa[1]*, Feng Cheng[1]*, Ying Zhu[1]*, Xuming Xu[1]*, Mingfeng Li[1]*, André M. M. Sousa[1,2], Mihovil Pletikos[1,3], Kyle A. Meyer[1], Goran Sedmak[1,3], Tobias Guennel[4], Yurae Shin[1], Matthew B. Johnson[1], Željka Krsnik[1], Simone Mayer[1,5], Sofia Fertuzinhos[1], Sheila Umlauf[6], Steven N. Lisgo[7], Alexander Vortmeyer[8], Daniel R. Weinberger[9], Shrikant Mane[6], Thomas M. Hyde[9,10], Anita Huttner[8], Mark Reimers[4], Joel E. Kleinman[9] & Nenad Šestan[1]

Brain development and function depend on the precise regulation of gene expression. However, our understanding of the complexity and dynamics of the transcriptome of the human brain is incomplete. Here we report the generation and analysis of exon-level transcriptome and associated genotyping data, representing males and females of different ethnicities, from multiple brain regions and neocortical areas of developing and adult post-mortem human brains. We found that 86 per cent of the genes analysed were expressed, and that 90 per cent of these were differentially regulated at the whole-transcript or exon level across brain regions and/or time. The majority of these spatio-temporal differences were detected before birth, with subsequent increases in the similarity among regional transcriptomes. The transcriptome is organized into distinct co-expression networks, and shows sex-biased gene expression and exon usage. We also profiled trajectories of genes associated with neurobiological categories and diseases, and identified associations between single nucleotide polymorphisms and gene expression. This study provides a comprehensive data set on the human brain transcriptome and insights into the transcriptional foundations of human neurodevelopment.

Human neurodevelopment is a complex and precisely regulated process that occurs over a protracted period of time[1–3]. Human-specific features of this process are likely to be important factors in the evolution of human specializations[2–5]. However, in addition to giving us remarkable cognitive and motor abilities, the formation of molecularly distinct and intricate neural circuits may have also increased our susceptibility to certain psychiatric and neurological disorders[4–9]. Furthermore, sex differences are important in brain development and function, and are a risk factor for conditions such as autism spectrum disorders (ASDs) and depression[9–13]. Research and progress in all these areas could be enhanced by a comprehensive analysis of the spatio-temporal dynamics of gene expression and transcript variants in the human brain.

Previous transcriptome studies of the developing human brain have used relatively small numbers of samples and predominantly focused on only a few regions or developmental time points[14–18]. In this Article, we explore the transcriptomes of 16 regions comprising the cerebellar cortex, mediodorsal nucleus of the thalamus, striatum, amygdala, hippocampus and 11 areas of the neocortex. The data set was generated from 1,340 tissue samples collected from 57 developing and adult post-mortem brains of clinically unremarkable donors representing males and females of multiple ethnicities.

Study design, data generation and quality control

To investigate the spatio-temporal dynamics of the human brain transcriptome, we created a 15-period system spanning the periods from embryonic development to late adulthood (Table 1 and Supplementary Information, section 2.1). We sampled transient prenatal structures and immature and mature forms of 16 brain regions, including 11 neocortex (NCX) areas, from multiple specimens per period (Table 2; Supplementary Information, section 2.2; Supplementary Figs 1–3; and Supplementary Table 1). The 11 NCX areas are collectively referred to hereafter as the region NCX. We also genotyped donor DNA using an Illumina 2.5-million single nucleotide polymorphism (SNP) chip (Supplementary Fig. 4 and Supplementary Table 2). Only brains from clinically unremarkable donors with no signs of large-scale genomic abnormalities were included in the study ($N = 57$, including 39 with both hemispheres; age, 5.7 weeks post-conception

Table 1 | Periods of human development and adulthood as defined in this study

Period	Description	Age
1	Embryonic	4 PCW ≤ Age < 8 PCW
2	Early fetal	8 PCW ≤ Age < 10 PCW
3	Early fetal	10 PCW ≤ Age < 13 PCW
4	Early mid-fetal	13 PCW ≤ Age < 16 PCW
5	Early mid-fetal	16 PCW ≤ Age < 19 PCW
6	Late mid-fetal	19 PCW ≤ Age < 24 PCW
7	Late fetal	24 PCW ≤ Age <38 PCW
8	Neonatal and early infancy	0 M (birth) ≤ Age <6 M
9	Late infancy	6 M ≤ Age < 12 M
10	Early childhood	1 Y ≤ Age <6 Y
11	Middle and late childhood	6 Y ≤ Age < 12 Y
12	Adolescence	12 Y ≤ Age < 20 Y
13	Young adulthood	20 Y ≤ Age < 40 Y
14	Middle adulthood	40 Y ≤ Age < 60 Y
15	Late adulthood	60 Y ≤ Age

M, postnatal months; PCW, post-conceptional weeks; Y, postnatal years.

[1]Department of Neurobiology and Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, Connecticut 06510, USA. [2]Graduate Program in Areas of Basic and Applied Biology, Abel Salazar Biomedical Sciences Institute, University of Porto, 4099-003 Porto, Portugal. [3]Graduate Program in Neuroscience, Croatian Institute for Brain Research, University of Zagreb School of Medicine, 10000 Zagreb, Croatia. [4]Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia 23298, USA. [5]MSc/PhD Molecular Biology Program, International Max Planck Research School for Molecular Biology, 37077 Göttingen, Germany. [6]Yale Center for Genome Analysis, Yale University School of Medicine, New Haven, Connecticut 06510, USA. [7]Institute of Genetic Medicine, Newcastle University, International Centre for Life, Newcastle upon Tyne NE1 3BZ, UK. [8]Department of Pathology, Yale University School of Medicine, New Haven, Connecticut 06510, USA. [9]Clinical Brain Disorders Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892, USA. [10]The Lieber Institute for Brain Development, Johns Hopkins University Medical Campus, Baltimore, Maryland 21205, USA.

*These authors contributed equally to this work.

©2011 Macmillan Publishers Limited. All rights reserved

Table 2 | Ontology and nomenclature of analysed brain regions and NCX areas

Periods 1 and 2	Periods 3–15
FC, frontal cerebral wall	OFC, orbital prefrontal cortex DFC, dorsolateral prefrontal cortex VFC, ventrolateral prefrontal cortex MFC, medial prefrontal cortex M1C, primary motor (M1) cortex
PC, parietal cerebral wall	S1C, primary somatosensory (S1) cortex IPC, posterior inferior parietal cortex
TC, temporal cerebral wall	A1C, primary auditory (A1) cortex STC, superior temporal cortex ITC, inferior temporal cortex
OC, occipital cerebral wall	V1C, primary visual (V1) cortex
HIP, hippocampal anlage	HIP, hippocampus
—	AMY, amygdala
VF, ventral forebrain MGE, medial ganglionic eminence LGE, lateral ganglionic eminence CGE, caudal ganglionic eminence	STR, striatum
DIE, diencephalon DTH, dorsal thalamus	MD, mediodorsal nucleus of the thalamus —
URL, upper (rostral) rhombic lip	CBC, cerebellar cortex

(PCW) to 82 years; sex, 31 males and 26 females; post-mortem interval, 12.11 ± 8.63 (mean ± s.d.) hours; pH, 6.45 ± 0.34 (mean ± s.d.)).

Transcriptome profiling was performed using total RNA extracted from a total of 1,340 dissected tissue samples (RNA integrity number, 8.83 ± 0.93 (mean ± s.d.); Supplementary Tables 3 and 4). We used the Affymetrix GeneChip Human Exon 1.0 ST Array platform, which features comprehensive coverage of the human genome, with 1.4 million probe sets that assay expression across the entire transcript or individual exon, thereby providing redundancy and increased confidence in estimates of gene-level differential expression (DEX, differentially expressed) and differential exon usage (DEU). Descriptions of tissue sampling and quality control measures implemented throughout transcriptome data generation steps are provided in Supplementary Information, sections 2–5, and Supplementary Figs 5–8.

Global transcriptome dynamics

Spatio-temporal gene expression

After quality control assessments and quantile normalization, we summarized core and unique probe sets, representing 17,565 mainly protein-coding genes, into gene-level information. Using stringent criteria ($\log_2$-transformed signal intensity of ≥6 in at least one sample and mean detection-above-background P value of <0.01 in at least one region of at least one period) to define an 'expressed' gene, we found that 15,132 (86.1%) of 17,565 genes surveyed were expressed in at least one brain region during at least one period, and that 14,375 (81.8%) were expressed in at least one NCX area (Fig. 1a, Supplementary Information 6.1 and Supplementary Fig. 9). To investigate the contributions of different factors to the global transcriptome dynamics, we applied multidimensional scaling and principal-component analysis, which revealed that region and age (that is, spatio-temporal dynamics) contribute more to the global differences in gene expression than do other tested variables: sex, ethnicity and inter-individual variation (Fig. 1b; Supplementary Information, sections 6.2 and 6.3; and Supplementary Figs 10 and 11).

To identify genes that were spatially or temporally regulated, we used a conservative threshold (false-discovery-rate Q value of <0.01 and ≥2-fold $\log_2$-transformed signal intensity difference), included post-mortem interval and RNA integrity number as technical covariates within all of our analysis-of-variance models of differential expression, considered the influence of dissection variation and applied a fivefold jackknife procedure (Supplementary Information, section 6.4, and Supplementary Figs 12 and 13). We found that 70.9% of expressed genes were spatially DEX between any two regions within at least one period, and that 24.1% were spatially DEX between any two NCX areas (Fig. 1a). By contrast, 89.9% of expressed genes were temporally DEX between any two periods across regions, and 85.3% were temporally DEX between any two periods across NCX areas. Moreover, 70.0% and 23.9% of expressed genes were both spatially and temporally DEX within brain regions and within NCX areas, respectively. The bulk of spatio-temporal regulation occurred during prenatal development. For instance, 57.7% of NCX-expressed genes were temporally DEX across fetal development (periods 3–7), whereas 9.1% were during postnatal development (periods 8–12) and 0.7% were during adulthood (periods 13–15). Together, these data indicate that the majority of brain-expressed protein-coding genes are temporally and, to a lesser extent, spatially regulated, and that this regulation occurs predominately during prenatal development.

Transcriptional architecture of the human brain

To assess transcriptional relatedness between brain regions/NCX areas, we calculated correlation matrices of pairwise comparisons

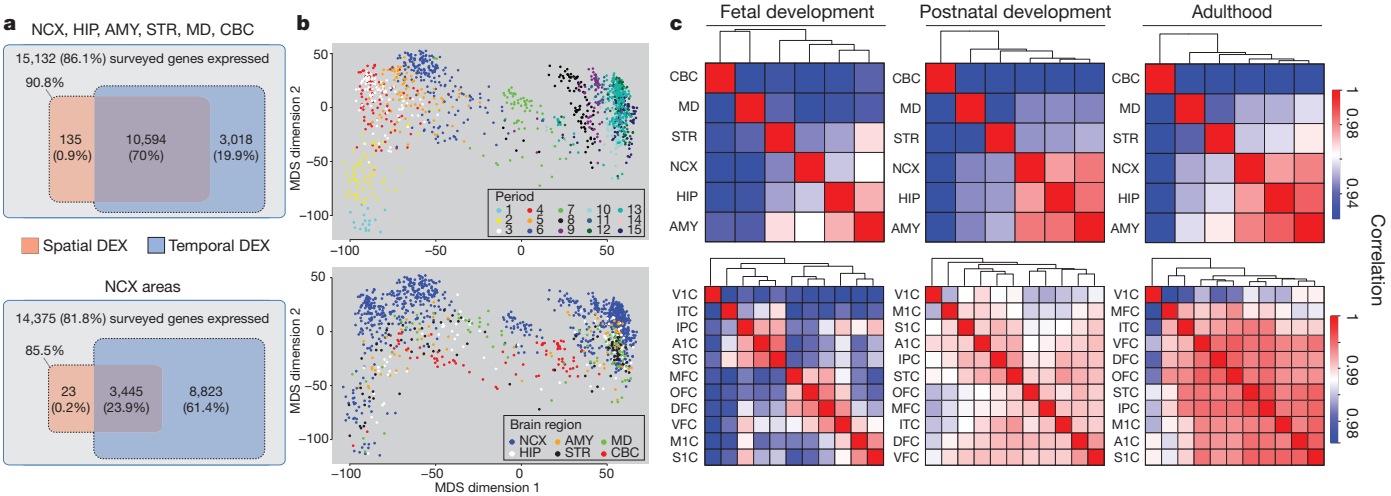


Figure 1 | Global spatio-temporal dynamics of gene expression. **a**, Venn diagrams representing the total number of genes considered to be expressed and the number of spatially and temporally DEX genes for brain regions (top) and NCX areas (bottom). **b**, Multidimensional scaling (MDS) plot showing transcriptional similarity, coloured by period (top) and region (bottom). Non-metric; stress = 18.9%. Euclidean distance of $\log_2$-transformed signal intensity was used to measure pairwise similarity. **c**, Heat map matrix of pairwise Spearman correlations between brain regions (top) and between NCX areas (bottom) during fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15).

©2011 Macmillan Publishers Limited. All rights reserved

(Fig. 1c) and performed unsupervised hierarchical clustering across periods 3–15, an interval during which all analysed regions/areas can be consistently followed across time. Among regions, this analysis showed distinct and developmentally regulated clustering of NCX (combination of 11 areas), HIP and AMY, with CBC having the most distinctive transcriptional profile. At the level of NCX areas, clustering formed the following groups during fetal periods: OFC, DFC and MFC; VFC and primary somatomotor cortex (S1C and M1C); and parietal–temporal perisylvian areas (IPC, A1C and STC). V1C had the most distinctive transcriptional profile of NCX areas throughout development and adulthood. The increased correlations between NCX, HIP, AMY and the majority of non-V1C NCX areas with age indicate that transcriptional differences are particularly pronounced during development.

Consistent with the clustering observed, CBC showed the greatest number of region-restricted or region-enriched DEX genes, with 516 (4.8%) of 10,729 genes spatially DEX (Supplementary Information, section 6.4, and Supplementary Table 5). By contrast, the numbers of genes highly enriched in the other regions were lower: NCX, 46 (0.43%); HIP, 48 (0.45%); AMY, 4 (0.04%); STR, 137 (1.28%); MD, 216 (2.01%). The majority of these spatially enriched genes were also temporally regulated, and some, such as those in Supplementary Figs 14 and 15 (NCX: *FLJ32063*, *KCNS1*; HIP: *CDC20B*, *METTL7B*; AMY: *TFAP2D*, *UTS2D*; STR: *C10orf11*, *PTPN7*; MD: *CEACAM21*, *SLC24A5*; CBC: *ESRRB*, *ZP2*), were transiently enriched during a narrow time window. These clustering and region-enrichment results reveal that regional transcriptomes are developmentally regulated and reflect anatomical differences.

Spatio-temporal differential exon usage

Alternative exon usage is an important mechanism for generating transcript diversity[19,20]. Using a splicing analysis of variance and a splicing index algorithm with conservative criteria ($Q < 0.01$ with a minimum twofold splice index difference between at least two regions/areas or periods; Supplementary Information, section 6.5), we found that 13,647 (90.2%) of 15,132 expressed genes showed DEU across sampled regions (0.1%), periods (19.5%) or both (70.6%). Of 14,375 NCX genes, 88.7% showed DEU across sampled areas (<0.01%), periods (59.8%) or both (28.9%). The regulation of DEU also varied in time, with the majority of expressed genes (83.0%) showing temporal DEU across fetal development, whereas only 0.9% and 1.4% were temporally regulated across postnatal development and adulthood, respectively.

Focusing on *ANKRD32*, a gene we have previously shown to express an alternative variant in the late mid-fetal frontal cortex[16], we confirmed and extended our findings on DEU by showing that whereas the longer isoform (*ANKRD32a*) was equally expressed across fetal NCX areas, the shorter isoform (*ANKRD32b*), comprising the last three exons, exhibited dynamic areal patterns. *ANKRD32b* was transiently expressed in a gradient along the anterior–posterior axis of the mid-fetal frontal cortex, with the highest expression in OFC and the lowest in M1C. Before this, *ANKRD32b* was most highly enriched in the ITC and, to a lesser extent, the STC. These spatio-temporal patterns disappeared after birth, when only *ANKRD32a* was expressed, and were not observed in the mouse NCX of equivalent ages (Supplementary Fig. 16 and Supplementary Table 6). These findings illustrate the complexity of DEU in the human brain and demonstrate how specific alternative transcripts can be spatially restricted during a narrow developmental window and with interspecies differences.

Sex differences in the transcriptome

Sex-biased gene expression

Previous studies have identified sexually dimorphic gene expression in the developing and adult human brain[11–13]. Analysis of our data set using a sliding-window algorithm and *t*-test model ($Q < 0.01$ with >2-fold difference in $\log_2$-transformed signal intensity; Supplementary Information, section 6.6) identified 159 genes, including a number of previously reported and newly uncovered genes with male or female bias in expression located on the Y (13 genes), X (9 genes) and autosomal (137 genes) chromosomes. A large fraction (76.7%) had male-biased expression (Fig. 2a and Supplementary Table 7). Notable spatial differences were observed, and more genes had sex-biased expression during prenatal development than during postnatal life, with the adult brain characterized by having the fewest.

Consistent with previous findings[12,13], we found that the largest differences were attributable to Y-chromosome genes, especially *PCDH11Y*, *RPS4Y1*, *USP9Y*, *DDX3Y*, *NLGN4Y*, *UTY*, *EIF1AY* and *ZFY*, which showed constant expression across regions and periods, with the exception of *PCDH11Y* downregulation in the postnatal CBC (Fig. 2b). Notably, the functional homologues of these genes on the X chromosome, barring *ZFX* during fetal development (*PCDH11X*, *RPS4X*, *USP9X*, *DDX3X*, *NLGN4X*, *UTX* and *EIF1AX*), were not upregulated in a compensatory manner in female brains (Supplementary Fig. 17).

We also found other X-linked and autosomal genes with sex-biased expression and distinct spatio-temporal patterns, including functionally uncharacterized transcripts (*LOC554203*, *C3orf62*, *FLJ35409* (also known as *MIR137HG*) and *DKFZP586I1420*), *S100A10* (which has been linked to depression[21]) and *IGF2* (an imprinted autosomal gene previously implicated in embryonic growth and cognitive function[22,23]), that showed population-level male-biased expression (Fig. 2c).

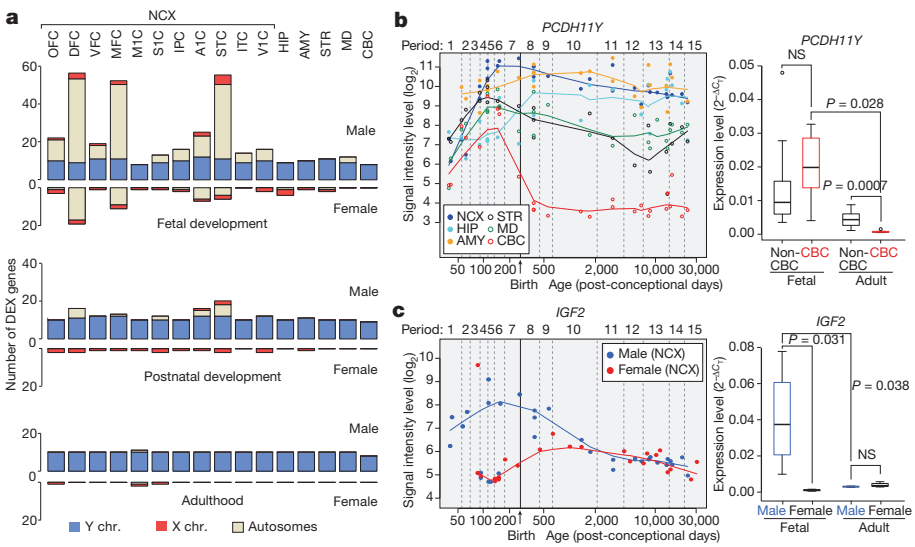


Figure 2 | Sex-biased gene expression.
a, Number of sex-biased DEX genes in brain regions/NCX areas during fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15).
b, *PCDH11Y* exon array signal intensity (left) and validation by quantitative PCR with reverse transcription (qRT–PCR; right) ($N = 5$ male brains per period). **c**, *IGF2* exon array signal intensity (left) and qRT–PCR (right) validation in NCX ($N = 4$ per sex and period). *P* values were calculated by unpaired *t*-test. Whiskers indicate fifth and ninety-fifth percentiles, respectively. NS, not significant.

©2011 Macmillan Publishers Limited. All rights reserved

Sex-biased exon usage

We next explored sex-biased DEU using a sliding-window algorithm with a splicing *t*-test model ($Q < 0.01$ and splicing index >2; Supplementary Information, section 6.6). We identified 155 genes (145 autosomal) that showed sex-biased expression of probe sets encoding one or a subset of exons (Supplementary Table 8) in one or multiple regions/NCX areas. These included several members of the collagen family of genes (*COL1A1*, *COL1A2*, *COL3A1*, *COL5A2* and *COL6A3*), *C3*, *KCNH2* (a gene associated with schizophrenia[24]), *NOTCH3* (a gene mutated in a common form of hereditary stroke disorder[25]), *ELN* (a gene located within the Williams syndrome critical region[26]) and *NLGN4X* (an X-chromosome gene implicated in synapse function and associated with ASD and moderate X-linked intellectual disability[10,27]). Although comparably expressed in males and females at the population and gene levels (Supplementary Fig. 17), *NLGN4X* had a significant male bias in expression of exon 7 and, to a lesser extent, exons 1, 5 and 6 in a developmentally regulated manner (Fig. 3). Together, these findings show that developmentally and spatially regulated differences in gene- and exon-level expression exist between male and female brains.

Gene co-expression networks

To extract additional biological information embedded in the multi-dimensional transcriptome data set, we performed weighted gene co-expression network analysis[28], which allowed us to identify modules of co-expressed genes. We identified 29 modules associated with distinct spatio-temporal expression patterns and biological processes (Fig. 4a; Supplementary Information, section 6.7; Supplementary Tables 9–11; and Supplementary Figs 18–20). Among modules corresponding to specific spatio-temporal patterns, M8 consisted of 24 genes with a common developmental trend that showed the highest expression levels in early fetal NCX and HIP (period 3), followed by a progressive decline in expression levels with age until infancy (period 9) (Fig. 4b). By contrast, M15 contained 310 genes showing changes in the opposite direction (relative to those in M8) in the NCX, HIP, AMY and STR (Fig. 4c). Gene ontology enrichment analysis showed that genes in M8 were enriched for gene ontology categories related to neuronal differentiation (Bonferroni-adjusted $P = 7.7 \times 10^{-3}$) and transcription factors ($P = 5.2 \times 10^{-3}$) (Supplementary Information, section 6.8, and Supplementary Table 9). Conversely, M15 gene ontology categories included ionic channels ($P = 8.0 \times 10^{-8}$) and neuroactive ligand–receptor interaction ($P = 4.0 \times 10^{-14}$).

Genes with the highest degree of connectivity within a module are termed hub genes and are expected to be functionally important within the module. M8 hub genes included transcription factors *TBR1*, *FEZF2*, *FOXG1*, *SATB2*, *NEUROD6* and *EMX1* (Fig. 4b), which have been functionally implicated in the development of NCX and HIP projection neurons[29–38]. Furthermore, *FOXG1* variants have been linked to Rett syndrome and intellectual disability[34]. Sequence variants in M15 hub genes (Fig. 4c) have been linked to major depression[39] (*GDA*) and to schizophrenia and affective disorders[6,40] (*NRGN* and *RGS4*).

We also identified two large-scale, temporally regulated modules (M20 and M2) with opposite developmental trajectories of genes co-expressed across regions: expression in M20 gradually decreased with age and expression in M2 gradually increased (Supplementary Figs 21 and 22). M20 was enriched for gene ontology categories related to zinc-finger proteins ($P = 7.3 \times 10^{-48}$) and transcription factors ($P = 4.8 \times 10^{-50}$), including many ZNF and SOX family members. M2 was enriched for gene ontology categories related to membrane proteins ($P = 1.8 \times 10^{-21}$), calcium signalling ($P = 8.1 \times 10^{-10}$), synaptic transmission ($P = 1.6 \times 10^{-6}$) and neuroactive ligand–receptor interaction ($P = 4.1 \times 10^{-4}$), reflecting processes important in postnatal brain maturation. Their hub genes encoded transcriptional factors, modulators of chromatin state and signal transduction proteins, all of which are likely to be involved in driving the co-expression networks. Drastic expression shifts in M20 and M2 in the opposite direction just before birth indicate that this period is associated with global transcriptional changes that probably reflect environmental influences on brain development and intrinsic changes in cellular composition and functional processes.

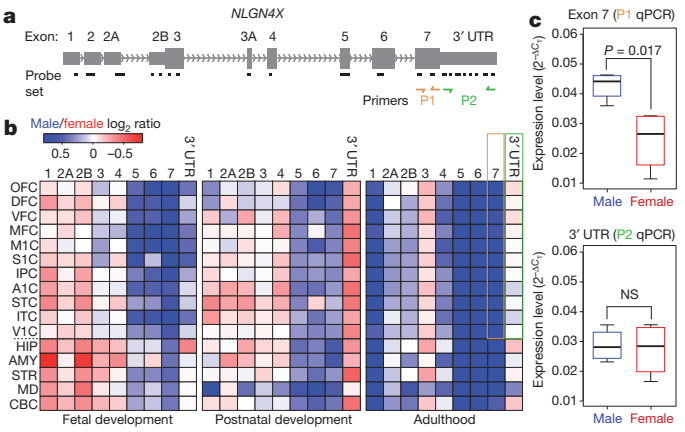


Figure 3 | Sex-biased differential exon usage. **a**, Gene structure and probe set composition of *NLGN4X*. Yellow and green arrows depict primers used for qRT–PCR validation. **b**, Heat map of the $\log_2$ male/female signal intensity ratio of each exon for fetal development (periods 3–7), postnatal development (periods 8–12) and adulthood (periods 13–15). Differences in expression of exon 7 (yellow frame) and the 3′ untranslated region (UTR; green frame) in adult NCX are highlighted. Note that exons 2 and 3A did not meet our expression criteria and are not represented. **c**, qRT–PCR validation of expression of exon 7 and the 3′ UTR in adult NCX ($N = 4$ per sex). *P* values were calculated by unpaired *t*-test. Whiskers indicate fifth and ninety-fifth percentiles, respectively.

Expression trajectories of neurodevelopment

One important use for the generated data set is to gain insight into normal and abnormal human neurodevelopment by analysing trajectories of individual genes or groups of genes associated with a particular neurobiological category or disease. To test this strategy, we compared our expression data for *DCX* (a gene expressed in neuronal progenitor cells and immature migrating neurons), as well as for genes associated with dendrite (*MAP1A*, *MAPT*, *CAMK2A*) and synapse (*SYP*, *SYPL1*, *SYPL2*, *SYN1*) development, with independently generated, non-transcriptome human data sets. The *DCX* expression trajectory was remarkably reminiscent of the reported changes in the density of DCX-immunopositive cells in the postnatal human HIP[36,40] ($r = 0.946$, Pearson correlation; Fig. 5a). In our transcriptome data set, *DCX* expression increased until early mid-fetal development (period 5) and then gradually declined with age until early childhood (period 10). Likewise, expression trajectories of dendrite and synapse development gene groups closely paralleled the growth of basal dendrites of DFC pyramidal neurons[41] ($r = 0.810$ for layer 3 and $r = 0.700$ for layer 5; Fig. 5b) and DFC synaptogenesis[42] ($r = 0.940$; Fig. 5c), respectively. Steep increases in both processes occurred between the late mid-fetal period and late infancy, indicating that a considerable portion of these two processes occurs before birth and reaches a plateau around late infancy.

After demonstrating the accuracy and viability of using the data set to profile human neurodevelopment, we manually curated lists of genes associated with over 80 categories, including various neurodevelopmental processes, neural cell types and neurotransmitter systems (Supplementary Information 6.9 and Supplementary Table 12). Notable trajectories and differences in their onset times, rates of increase and decrease, and shapes were observed within and between brain regions for categories including major neurodevelopmental processes (neural cell proliferation and migration, dendrite and synapse development, and myelination; Fig. 5d), cortical GABAergic inhibitory interneurons (*CALB1*, *CALB2*, *NOS1*, *PVALB* and *VIP*) and

©2011 Macmillan Publishers Limited. All rights reserved

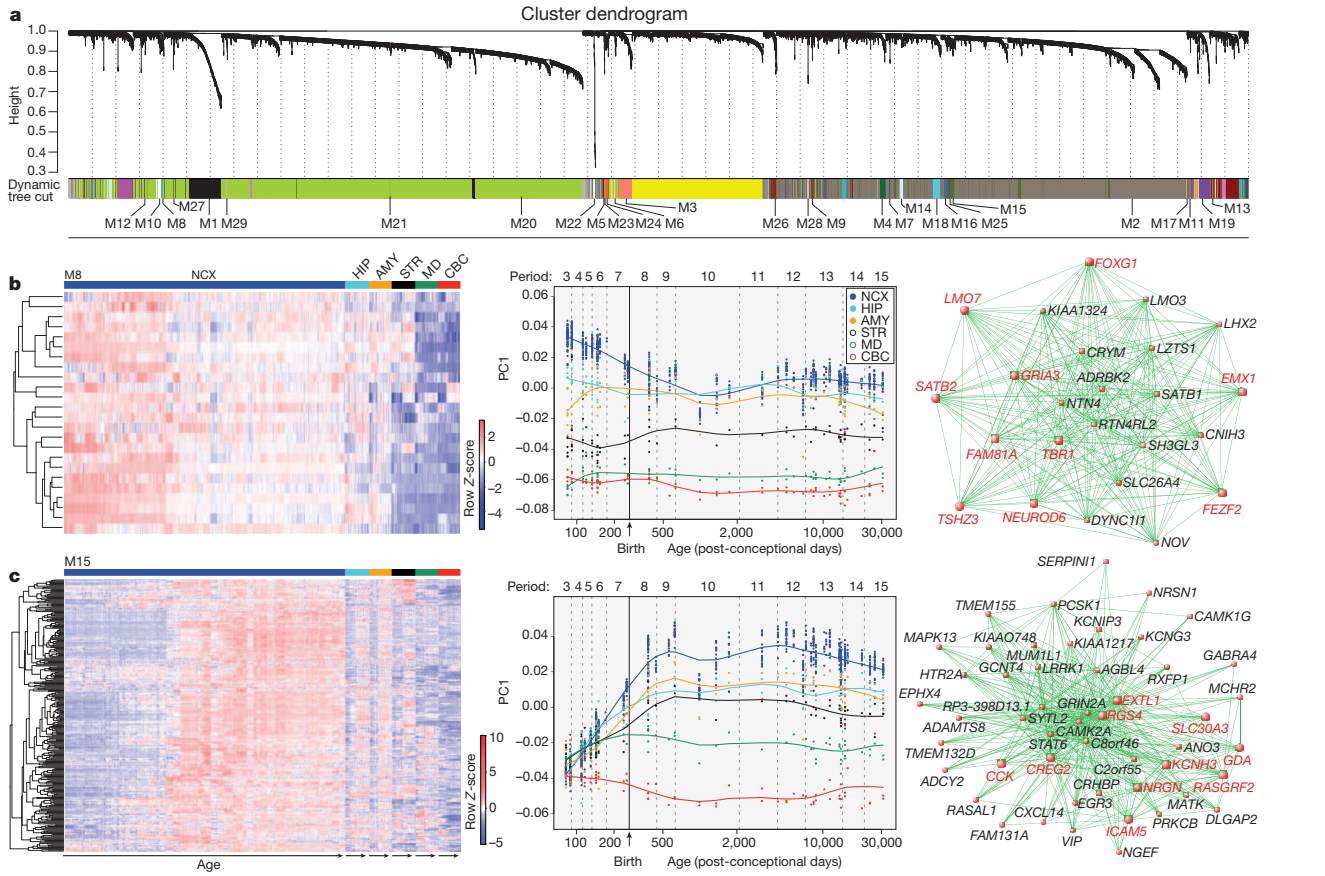


Figure 4 | Global co-expression networks and gene modules. a, Dendrogram from gene co-expression network analysis of samples from periods 3–15. Modules of co-expressed genes were assigned a colour and number (M1 to M29). **b**, Left: heat map of genes in M8 showing the spatio-temporal expression pattern after hierarchical clustering. The expression values for each gene are arranged in the heat map, ordered first by brain region, then by age and last by NCX area. Middle: spatio-temporal pattern of M8 summarized by the first principal component (PC1) for expression of genes in the module across age. Right: 24 M8 genes; the top ten hub genes are shown in red. **c**, Same analyses as in **b**, but for M15; the top 50 genes defined by the highest intramodular connectivity are shown in right panel. Results for other modules are available in Supplementary Information.

glutamate receptors (Supplementary Figs 23 and 24). Two expected patterns were observed in neurodevelopmental trajectories: changes in expression of cell proliferation genes preceded the increase in expression of *DCX*, and expression of each decreased during perinatal development whereas synapse development, dendrite development and myelination trajectories increased. Notably, the NCX trajectory for synapse development did not drastically decline during late childhood or adolescence (Fig. 5c, d) as previously reported for synapse density[42]. We also identified co-expression network modules and additional genes that are highly correlated with the categories (Supplementary Tables 10, 13 and 14). For example, M20 and M2 were strongly correlated with neuron migration ($r = 0.894$) and myelination ($r = 0.972$), respectively.

In addition, our data set enabled us to generate expression trajectories of genes commonly associated with ASD and schizophrenia. We investigated a number of genes previously linked to these disorders (Supplementary Information, section 6.10) and observed distinct and dynamic expression patterns, especially among NCX areas (Supplementary Fig. 25 shows examples for *CNTNAP2*, *MET*, *NLGN4X* and *NRGN*). To gain insight into potential biological functions of ASD- and schizophrenia-associated genes in human neurodevelopment, we identified other genes with significantly correlated spatio-temporal expression profiles and performed gene ontology enrichment analysis (Supplementary Tables 15 and 16). These findings reveal associated spatio-temporal differences in these expression trajectories and provide additional co-expressed genes that can be interrogated for their role in the respective processes or disorders.

Expression quantitative trait loci

Previous studies have identified expression quantitative trait loci (eQTLs) in the adult human brain, primarily in the cerebral cortex[43–47]. Our multiregional developmental data set enabled us to search for association between SNP genotypes and spatio-temporal gene expression. We tested only for *cis*-eQTLs, restricting the search to SNPs within 10 kilobases of either a transcription start site or a transcription end site, as opposed to *trans*-eQTLs, which would require much larger sample sizes.

Implementing a conservative strategy (gene-wide Bonferroni correction followed by genome-wide $Q < 0.1$; Supplementary Information, section 9), we identified 39 NCX, eight HIP, four AMY, two STR, six MD and five CBC genes (Supplementary Table 17) with evidence of *cis*-eQTL, including two previously reported genes[45,47] (*ITGB3BP* and *ANKRD27*). Consistent with previous studies[48], associated SNPs were enriched near transcription start and termination sites (Fig. 6a, b).

An example of a significant association in NCX, MD and CBC is that between SNP rs10785190 and *GLIPR1L2*, a member of the glioma pathogenesis-related 1 family of genes[49]. The expression differences were observed at the level of the whole transcript and exons 1 and 2, the only exons we observed to be expressed at appreciable levels in the NCX (Fig. 6c, d). The NCX probably had more *cis*-eQTLs than other regions owing to its smaller variation in gene expression resulting from the averaged expression of 11 areas. Many eQTLs identified as significant in NCX also have similar associations in other regions, although they were not statistically significant after the conservative genome-wide correction (Supplementary Table 17). Thus, we have

©2011 Macmillan Publishers Limited. All rights reserved

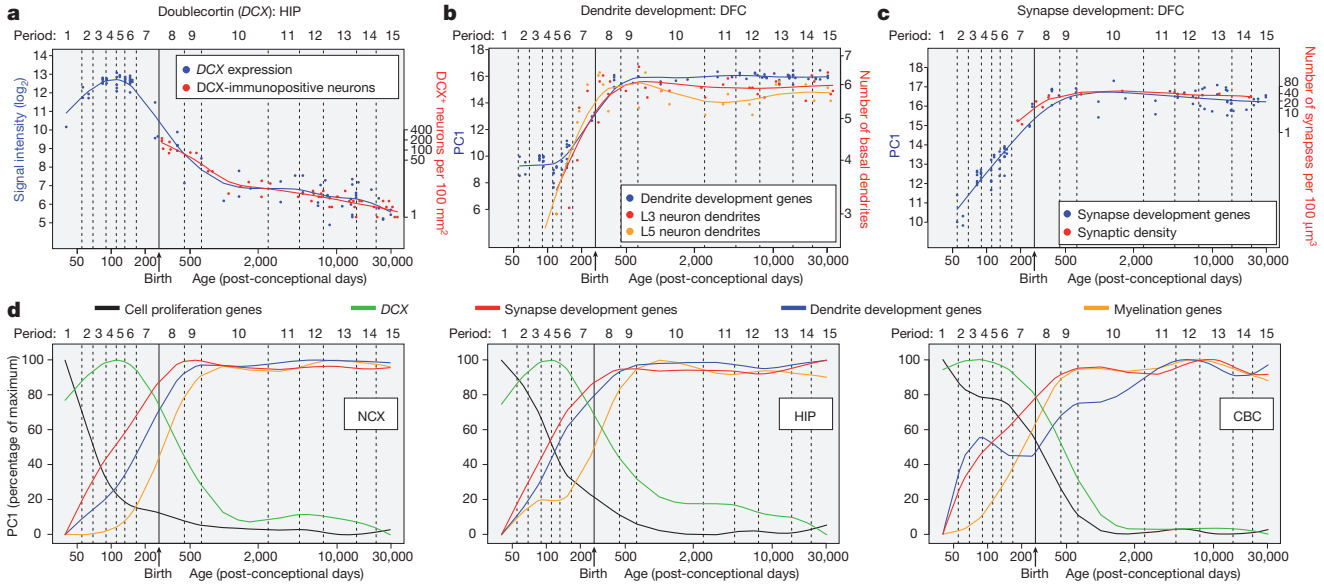


Figure 5 | Trajectories of genes associated with neurodevelopmental processes. **a**, Comparison between *DCX* expression in HIP and the density of DCX-immunopositive cells in the human dentate gyrus[36]. **b**, Comparison between transcriptome-based dendrite development trajectory in DFC and Golgi-method-based growth of basal dendrites of layer 3 (L3) and 5 (L5) pyramidal neurons in the human DFC[41]. **c**, Comparison between transcriptome-based synapse development trajectory in DFC and density of DFC synapses calculated using electron microscopy[42]. For **b** and **c**, PC1 for gene expression was plotted against age to represent the developmental trajectory of genes associated with dendrite (**b**) or synapse (**c**) development. Independent data sets were centred, scaled and plotted on a logarithmic scale. **d**, PC1 value for the indicated sets of genes (expressed as percentage of maximum) plotted against age to represent general trends and regional differences in several neurodevelopmental processes in NCX, HIP and CBC.

identified polymorphic regulators of transcription in different regions across development, potentially providing insights into inter-individual differences and genetic control of the brain transcriptome.

Discussion

Our analysis reveals several features of the human brain transcriptome, and increases our knowledge of the transcriptional events in human neurodevelopment. We show that gene expression and exon usage have complex and dynamically regulated patterns, some of which may not be evident in the transcriptomes of commonly studied model organisms. Moreover, these patterns differ more prominently across time and space than they do between sexes, ethnicities or individuals, despite their underlying genetic differences. Transcriptome differences between males and females also included several disease-related genes, offering possible mechanisms underlying the sex differences in the incidence, prevalence and severity of some brain disorders. We also found that some of the inter-individual variations in the regional and developmental transcriptomes were associated with specific SNP genotypes, which may have altered expression-regulating elements. Thus, the present data set (available at http://www.humanbraintranscriptome.org), along with an accompanying study[50], provides a basis for a variety of further investigations and comparisons with other transcriptome-related data sets of both healthy and diseased states.

Although our study has uncovered many intricacies in gene expression and exon usage in the human brain, there are potential limitations of our study that warrant discussion. Foremost, we used stringent criteria to minimize false positives and faithfully characterize general transcriptional patterns, rather than to capture all the changes that may occur. Also, we analysed dissected tissue that contained multiple cell types, thus diluting the transcriptional contribution and dynamic range of expression of any one specific cell type. Current limitations prevent us from using cell-type-specific approaches in systematically analysing the spatio-temporal transcriptome. Furthermore, the number of brains and regions analysed so far is not sufficient to investigate the full magnitude of transcriptional changes or the full range of eQTLs. Application of sequencing technology will allow more in-depth analyses of the transcriptome, and aid in discovery of novel or low-expressing transcripts. Finally, although specific patterns of expression are often linked to specialized biological processes, it is important to remember that the relationship between messenger RNA and protein levels is not always linear nor translated into apparent phenotypic differences. As these concerns are addressed in future

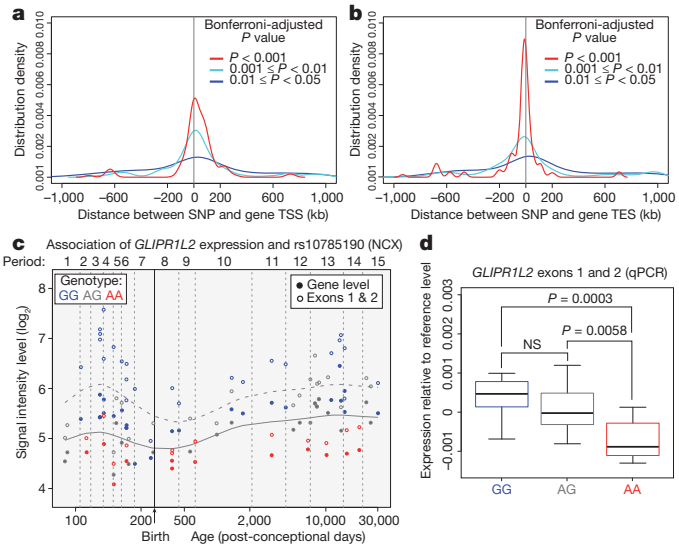


Figure 6 | Association between SNPs and gene expression. **a**, **b**, SNP distribution around transcription start sites (TSS; **a**) and transcription end sites (TES; **b**) of the associated genes, based on several Wald test *P*-value cut-offs after gene-wide Bonferroni correction. **c**, *GLIPR1L2* expression association with rs10785190 genotype, a SNP located in exon 1. The solid and dashed curves, calculated from locally weighted scatter-plot smoothing (LOWESS), show the developmental trends of gene expression and exon-1 and exon-2 expression, respectively. **d**, qRT–PCR validation of exon-1 and exon-2 expression in NCX for each genotype ($N = 14$ GG, 14 AG, 8 AA), plotted relative to the LOWESS curve in **c** to facilitate comparison across developmental periods. *P* values were calculated by unpaired *t*-test. Whiskers indicate fifth and ninety-fifth percentiles, respectively.

©2011 Macmillan Publishers Limited. All rights reserved

with more samples and new data sets from human and non-human primate brains, it will be possible to uncover deeper insights into the transcriptional foundations of human brain development and evolution.

METHODS SUMMARY

Supplementary Information, sections 3–9, provides a full description of tissue acquisition and processing, data generation, validation and analyses.

Received 11 December 2010; accepted 30 August 2011.

1. Kostovic, I. & Judas, M. Prolonged coexistence of transient and permanent circuitry elements in the developing cerebral cortex of fetuses and preterm infants. *Dev. Med. Child Neurol.* **48,** 388–393 (2006).
2. Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nature Rev. Neurosci.* **10,** 724–735 (2009).
3. Rubenstein, J. L. Annual Research Review: Development of the cerebral cortex: implications for neurodevelopmental disorders. *J. Child Psychol. Psychiatry* **52,** 339–355 (2011).
4. Preuss, T., Cáceres, M., Oldham, M. & Geschwind, D. Human brain evolution: insights from microarrays. *Nature Rev. Genet.* **5,** 850–860 (2004).
5. Hill, R. S. & Walsh, C. A. Molecular insights into human brain evolution. *Nature* **437,** 64–67 (2005).
6. Lewis, D. A. & Levitt, P. Schizophrenia as a disorder of neurodevelopment. *Annu. Rev. Neurosci.* **25,** 409–432 (2002).
7. Meyer-Lindenberg, A. & Weinberger, D. R. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Rev. Neurosci.* **7,** 818–827 (2006).
8. Insel, T. Rethinking schizophrenia. *Nature* **468,** 187–193 (2010).
9. State, M. The genetics of child psychiatric disorders: focus on autism and Tourette syndrome. *Neuron* **68,** 254–269 (2010).
10. Jamain, S. *et al.* Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nature Genet.* **34,** 27–29 (2003).
11. Vawter, M. *et al.* Gender-specific gene expression in post-mortem human brain: localization to sex chromosomes. *Neuropsychopharmacology* **29,** 373–384 (2004).
12. Weickert, C. *et al.* Transcriptome analysis of male-female differences in prefrontal cortical development. *Mol. Psychiatry* **14,** 558–561 (2009).
13. Reinius, B. & Jazin, E. Prenatal sex differences in the human brain. *Mol. Psychiatry* **14,** 988–989 (2009).
14. Abrahams, B. *et al.* Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc. Natl Acad. Sci. USA* **104,** 17849–17854 (2007).
15. Sun, T. *et al.* Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science* **308,** 1794–1798 (2005).
16. Johnson, M. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62,** 494–509 (2009).
17. Somel, M. *et al.* MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.* **20,** 1207–1218 (2010).
18. Ip, B. *et al.* Investigating gradients of gene expression involved in early human cortical development. *J. Anat.* **217,** 300–311 (2010).
19. Licatalosi, D. D. & Darnell, R. B. Splicing regulation in neurologic disease. *Neuron* **52,** 93–101 (2006).
20. Blencowe, B. J. Alternative splicing: new insights from global analyses. *Cell* **126,** 37–47 (2006).
21. Svenningsson, P. *et al.* Alterations in 5–HT1B receptor function by p11 in depression-like states. *Science* **311,** 77–80 (2006).
22. Chen, D. Y. *et al.* A critical role for IGF-II in memory consolidation and enhancement. *Nature* **469,** 491–497 (2011).
23. Lehtinen, M. K. *et al.* The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. *Neuron* **69,** 893–905 (2011).
24. Huffaker, S. J. *et al.* A primate-specific, brain isoform of KCNH2 affects cortical physiology, cognition, neuronal repolarization and risk of schizophrenia. *Nature Med.* **15,** 509–518 (2009).
25. Joutel, A. *et al.* Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* **383,** 707–710 (1996).
26. Ewart, A. K. *et al.* Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. *Nature Genet.* **5,** 11–16 (1993).
27. Südhof, T. C. Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455,** 903–911 (2008).
28. Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4,** article17 (2005).
29. Hansen, D. V., Rubenstein, J. L. & Kriegstein, A. R. Deriving excitatory neurons of the neocortex from pluripotent stem cells. *Neuron* **70,** 645–660 (2011).
30. Hevner, R. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29,** 353–366 (2001).
31. Molyneaux, B. J., Arlotta, P., Hirata, T., Hibi, M. & Macklis, J. D. Fezl is required for the birth and specification of corticospinal motor neurons. *Neuron* **47,** 817–831 (2005).
32. Chen, B., Schaevitz, L. & McConnell, S. Fezl regulates the differentiation and axon targeting of layer 5 subcortical projection neurons in cerebral cortex. *Proc. Natl Acad. Sci. USA* **102,** 17184–17189 (2005).
33. Chen, J., Rasin, M., Kwan, K. & Sestan, N. Zfp312 is required for subcortical axonal projections and dendritic morphology of deep-layer pyramidal neurons of the cerebral cortex. *Proc. Natl Acad. Sci. USA* **102,** 17792–17797 (2005).
34. Ariani, F. *et al.* FOXG1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet.* **83,** 89–93 (2008).
35. Kwan, K. *et al.* SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. *Proc. Natl Acad. Sci. USA* **105,** 16021–16026 (2008).
36. Knoth, R. *et al.* Murine features of neurogenesis in the human hippocampus across the lifespan from 0 to 100 years. *PLoS ONE* **5,** e8809 (2010).
37. Han, W. *et al.* TBR1 directly represses Fezf2 to control the laminar origin and development of the corticospinal tract. *Proc. Natl Acad. Sci. USA* **108,** 3041–3046 (2011).
38. McKenna, W. L. *et al.* Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development. *J. Neurosci.* **31,** 549–564 (2011).
39. Perroud, N. *et al.* Genome-wide association study of increasing suicidal ideation during antidepressant treatment in the GENDEP project. *Pharmacogenomics J.* advance online publication, ⟨http://dx.doi.org/10.1038/tpj.2010.70⟩ (2010).
40. Stefansson, H. *et al.* Common variants conferring risk of schizophrenia. *Nature* **460,** 744–747 (2009).
41. Petanjek, Z., Judas, M., Kostović, I. & Uylings, H. Lifespan alterations of basal dendritic trees of pyramidal neurons in the human prefrontal cortex: a layer-specific pattern. *Cereb. Cortex* **18,** 915–929 (2008).
42. Huttenlocher, P. R. & Dabholkar, A. S. Regional differences in synaptogenesis in human cerebral cortex. *J. Comp. Neurol.* **387,** 167–178 (1997).
43. Stranger, B. E. *et al.* Population genomics of human gene expression. *Nature Genet.* **39,** 1217–1224 (2007).
44. Heinzen, E. L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol.* **6,** e1 (2008).
45. Liu, C. *et al.* Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol. Psychiatry* **15,** 779–784 (2010).
46. Gibbs, J. R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6,** e1000952 (2010).
47. Myers, A. J. *et al.* A survey of genetic human cortical gene expression. *Nature Genet.* **39,** 1494–1499 (2007).
48. Webster, J. A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.* **84,** 445–458 (2009).
49. Ren, C., Ren, C. H., Li, L., Goltsov, A. A. & Thompson, T. C. Identification and characterization of RTVP1/GLIPR1-like genes, a novel p53 target gene cluster. *Genomics* **88,** 163–172 (2006).
50. Colantuoni, C. *et al.* Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* doi:10.1038/nature10524 (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Belanger, V. Imamovic, R. Johnson, P. Larton, S. Lindsay, B. Poulos, J. Rajan, D. Rimm and R. Zielke for assistance with tissue acquisition, D. Singh for technical assistance, I. Kostovic and Z. Petanjek for dendrite measurements, P. Levitt for suggesting the inclusion of ITC in the study, and D. Karolchik and A. Zweig for help in creating tracks for the UCSC Genome Browser. We also thank A. Beckel-Mitchener, M. Freund, M. Gerstein, D. Geschwind, T. Insel, M. Judas, J. Knowles, E. Lein, P. Levitt, N. Parikshak and members of the Sestan laboratory for discussions and criticism. Tissue was obtained from several sources including the Human Fetal Tissue Repository at the Albert Einstein College of Medicine, the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, the Laboratory of Developmental Biology at the University of Washington (supported by grant HD000836 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development) and the Joint MRC/Wellcome Trust Human Developmental Biology Resource (http://hdbr.org) at the IHG, Newcastle Upon Tyne (UK funding awards G0700089 and GR082557). Support for predoctoral fellowships was provided by the China Scholarship Council (Y.Z.), the Portuguese Foundation for Science and Technology (A.M.M.S.), the Samsung Scholarship Foundation (Y.S.), a Fellowship of the German Academic Exchange Service – DAAD (S. Mayer) and NIDA grant DA026119 (T.G.). This work was supported by grants from the US National Institutes of Health (MH081896, MH089929, NS054273), the Kavli Foundation and NARSAD, and by a James S. McDonnell Foundation Scholar Award (N.S.).

Author Contributions H.J.K., Y.I.K., A.M.M.S., M.P., K.A.M., G.S., Y.S., M.B.J., Z.K., S. Mayer, S.F., S.U., S. Mane and N.S. performed and analysed the experiments. F.C., Y.Z., X.X., M.L., T.G. and M.R. analysed the data. Z.K., A.M.M.S., M.P., G.S., S.N.L, A.V., D.R.W., T.M.H., A.H., J.E.K. and N.S. participated in tissue procurement and examination. N.S. designed the study and wrote the manuscript, which all authors commented and edited.

Author Information Exon array have been deposited in the NCBI Gene Expression Omnibus under accession number GSE25219 and genotyping data have been deposited in the NCBI database of Genotypes and Phenotypes under accession number phs000406.v1.p1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to N.S. (nenad.sestan@yale.edu).

©2011 Macmillan Publishers Limited. All rights reserved